J. Appl. Math. & Informatics Vol. 29(2011), No. 5 - 6, pp. 1435 - 1451
Website: http://www.kcam.biz

RELIABILITY ANALYSIS FOR THE TWO-PARAMETER PARETO DISTRIBUTION UNDER RECORD VALUES

LIANG WANG*, YIMIN SHI, PING CHANG

ABSTRACT. In this paper the estimation of the parameters as well as survival and hazard functions are presented for the two-parameter Pareto distribution by using Bayesian and non-Bayesian approaches under upper record values. Maximum likelihood estimation (MLE) and interval estimation are derived for the parameters. Bayes estimators of reliability performances are obtained under symmetric (Squared error) and asymmetric (Linex and general entropy (GE)) losses, when two parameters have discrete and continuous priors, respectively. Finally, two numerical examples with real data set and simulated data, are presented to illustrate the proposed method. An algorithm is introduced to generate records data, then a simulation study is performed and different estimates results are compared.

AMS Mathematics Subject Classification : 62F10, 62F15.
Key words and phrases : Reliability, two-Parameter Pareto distribution, Bayes analysis, record values, loss function, Monte-Carlo simulation.

1. Introduction

Let $\{X_n, n = 1, 2, \cdots\}$ be a sequence of independent and identically distributed (i.i.d.) random variables with a cumulative distribution function (cdf) $F(x)$ and probability density function (pdf) $f(x)$. An observation X_j is called an upper record value if its value exceeds that of all previous observations. Thus X_j is an upper record value if $X_j > X_i$ for each $i < j$. The record times sequence $\{T_n, n \geq 0\}$ is defined in the following manner

$$T_0 = 1 \text{ with probability } 1$$

and for $n \geq 1$,

$$T_n = \min\{j : X_j > X_{T_{n-1}}\}.$$

Received January 20, 2011. Revised March 28, 2011. Accepted April 4, 2011. *Corresponding author.

Furthermore, the sequence $\{X_{T_n}, n = 0, 1, 2, \cdots\}$ is called the sequence of upper records of the original sequence. An analogous definition can be given for lower record values when the sign $' >'$ is changed into $' <'$. In this paper, we only concerned with the upper record values, and similar results can be derived for lower record values.

Chandler [7] introduced the study of record values and documented many of the basic properties of records, which is a special order statistic from a sample whose size is determined by the values and the order of occurrence of observations. Record values are of interest and of importance in many real life applications, such as weather, sports, economics, life-tests, stock market and so on. A growing interest in records has arisen in the last two decades and their properties have been extensively studied in literature. See for instance, Al-Hussaini and Ahmed [3] and Mohamed and Mohamed [12]. For applications of the record values see Nevzorov [13]. More details about record values can be found in, i.e. Ahsanullah [1] and Arnold et al. [4].

For the sake of simplicity, let $X = (X_{T_1}, X_{T_2}, \ldots, X_{T_n}) = (X_1, X_2, \ldots, X_n)$ be upper record values from pdf $f(x)$ and cdf $F(x)$, $x = (x_1, x_2, \ldots, x_n)$ is an observe value of X, then the joint pdf of X, (see Ahsanullah [1], Arnold et al. [4]), can be expressed as

$$f(x) = f(x_n)\prod_{i=1}^{n-1}\frac{f(x_i)}{1-F(x_i)}. \tag{1}$$

Consider two-parameter Pareto distribution with cdf and pdf as follows

$$f(x;\theta,\beta) = \beta\theta^{\beta}x^{-(\beta+1)}, F(x;\theta,\beta) = 1-\theta^{\beta}x^{-\beta}, \beta > 0, x \geq \theta > 0, \tag{2}$$

where β and θ are unknown parameters.

The survival function $R(t)$ and the hazard function $H(t)$ of (2) at mission time t are given by

$$R(t) = 1 - F(t;\theta,\beta) = \theta^{\beta}t^{-\beta},\ H(t) = \frac{f(t;\theta,\beta)}{1-F(t;\theta,\beta)} = t^{-1}\beta. \tag{3}$$

The Pareto distribution was introduced by Pareto [14] as a model used to describe income distributions as well as a wide variety of other socio-economic phenomena such as insurance claims, firm assets, stock price fluctuations and the occurrence of natural phenomena. As a useful model, however, Pareto distribution also has many applications in life test and reliability studies. There are also several authors who discussed the Pareto distribution as a lifetime model. For instance, Tiwari et el. [19] discussed the estimation of reliability performances by using a fully Bayes approach wherein the range depends on one of the parameters. Wu et al. [21] considered the weighted moments estimators of scale parameter θ of Pareto distribution with known shape β parameter based on multiply type-II censored sample of electronic components. Dayna et al. [8]

discussed a goodness of fit test problem for the Pareto distribution under observations are type-II right censored data. For more details about the Pareto distribution as a lifetime model we refer to Johnson et al. [10] and Soliman [16].

Suppose that $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$ are n observed upper record values from the two-parameter Pareto distribution (2) with the parameters β and θ, then by (1) the joint pdf of the records sample $X = (X_1, X_2, \ldots, X_n)$ is given by

$$f_X(x; \beta, \theta) = \beta^n \theta^\beta x_n^{-\beta} \prod_{i=1}^{n} \frac{1}{x_i}. \tag{4}$$

The organization of this paper is as follows. Section 2 is devoted to give some frequentist estimation for parameters β and θ, such as maximum likelihood estimation (MLE) and interval estimation. Bayes analysis is presented in Section 3 for reliability performances under symmetric and asymmetric loss functions. Two numerical examples and some conclusions are given in Section 4 and Section 5, respectively.

2. Frequentist estimation

In this section we consider the frequentist estimation for the Pareto parameters, which include MLEs, as well as interval estimation for β and θ.

2.1. Point estimation. From (4), the log-likelihood function can be expressed as

$$L(\theta, \beta) \quad \propto \quad n \ln \beta + \beta \ln \theta - \beta \ln x_n. \tag{5}$$

Since function $L(\theta, \beta)$ is increasing in θ, the MLE for θ, namely $\hat{\theta}_M$, is given by

$$\hat{\theta}_M = X_1. \tag{6}$$

Substituting the MLE of θ into (5), we can get the derivative function in β as follows

$$\frac{\partial}{\partial \beta} L(\theta, \beta) = \frac{n}{\beta} + \ln x_1 - \ln x_n,$$

then the MLE of β, namely $\hat{\beta}_M$, can be written as

$$\hat{\beta}_M = n \left[\ln X_n - \ln X_1\right]^{-1}. \tag{7}$$

2.2. Approximate interval estimation. From the log-likelihood function (5), we have

$$\frac{\partial^2 L(\beta, \theta)}{\partial \beta^2} = -\frac{n}{\beta}, \frac{\partial^2 L(\beta, \theta)}{\partial \theta^2} = -\frac{\beta}{\theta^2}, \tag{8}$$

and

$$\frac{\partial^2 L(\beta, \theta)}{\partial \beta \partial \theta} = \frac{1}{\theta}. \tag{9}$$

The Fisher information matrix $I(\beta,\theta)$ is derived by taking expectations of minus Eqs. (8) and (9). Under some mild regular conditions, $(\hat{\beta}_M,\hat{\theta}_M)$ is approximately brivariately normal with mean (β,θ) and covariance matrix $I^{-1}(\beta,\theta)$. In application, we use $I^{-1}(\hat{\beta}_M,\hat{\theta}_M)$ to estimate $I^{-1}(\beta,\theta)$, then the approximate distribution of (β,θ) can be expressed as

$$\left(\hat{\beta}_M,\hat{\theta}_M\right)\sim\ N\left((\beta,\theta),I_0^{-1}(\hat{\beta}_M,\hat{\theta}_M)\right)_{(\hat{\beta}_M,\hat{\theta}_M)}, \tag{10}$$

where

$$I_0(\hat{\beta}_M,\hat{\theta}_M)=\begin{bmatrix} -\frac{\partial^2 L(\beta,\theta)}{\partial\beta^2} & -\frac{\partial^2 L(\beta,\theta)}{\partial\beta\partial\theta} \\ -\frac{\partial^2 L(\beta,\theta)}{\partial\beta\partial\theta} & -\frac{\partial^2 L(\beta,\theta)}{\partial\theta^2}\end{bmatrix}_{(\hat{\beta}_M,\hat{\theta}_M)}.$$

From the approximate distribution of (β,θ), for any $0<\alpha<1$, a $100(1-\alpha)\%$ approximate confidence intervals for β and θ are given by

$$\left(\hat{\beta}_M-Z_{\frac{\alpha}{2}}\sqrt{V_{11}},\hat{\beta}_M+Z_{\frac{\alpha}{2}}\sqrt{V_{11}}\right)\text{ and }\left(\hat{\theta}_M-Z_{\frac{\alpha}{2}}\sqrt{V_{22}},\hat{\theta}_M+Z_{\frac{\alpha}{2}}\sqrt{V_{22}}\right), \tag{11}$$

respectively, where V_{11} and V_{22} are the elements on the main diagonal of the covariance matrix $I_0^{-1}(\hat{\beta}_M,\hat{\theta}_M)$ and $Z_{\frac{\alpha}{2}}$ is the $\frac{\alpha}{2}$ right-tail percentile of the standard normal distribution.

Furthermore, the normal approximation for $(\hat{\beta}_M,\hat{\theta}_M)$ implies that the statistic

$$\left[\hat{\beta}_M-\beta\hat{\theta}_M-\theta\right]I_0^{-1}(\hat{\beta}_M,\hat{\theta}_M)\left[\hat{\beta}_M-\beta\hat{\theta}_M-\theta\right]'$$

has an asymptotical chi-squared distribution with two degrees of freedom. Furthermore, a $100(1-\alpha)\%$ approximate confidence region for (β,θ) can be expressed as

$$\left\{(\beta,\theta):\left[\hat{\beta}_M-\beta\hat{\theta}_M-\theta\right]I_0^{-1}(\hat{\beta}_M,\hat{\theta}_M)\left[\hat{\beta}_M-\beta\hat{\theta}_M-\theta\right]'\le\chi_2^2(\alpha)\right\}, \tag{12}$$

where $\chi_a^2(p)$ is the $100p\%$ right-tail percentile of chi-squared distribution with a degrees of freedom.

2.3. Exact interval estimation. Denote $Y_i=-\beta(\ln\theta-\ln\ X_i)$, it can be seen that $Y_i, i=1,2,\ldots,n$, are record values from standard exponential distribution with mean 1. Since the exponential distribution has the lack of memory property and consequently the differences between successive records will be i.i.d. samples from standard exponential distribution.

Denote

$$\begin{aligned} Z_1 &= Y_1=\beta[\ln X_1-\ln\theta], \\ Z_2 &= Y_2-Y_1=\beta[\ln X_2-\ln X_1], \\ &\cdots \\ Z_n &= Y_n-Y_{n-1}=\beta[\ln X_n-\ln X_{n-1}]. \end{aligned}$$

It is noted that $Z_i, i = 1, 2, \ldots, n$, are independent and identical distributed as standard exponential distribution with mean 1. Hence

$$\kappa = 2Z_1 = 2\beta[\ln X_1 - \ln\theta]$$

has a chi-squared distribution with 2 degrees of freedom and

$$\varepsilon = 2\sum_{i=2}^{n} Z_i = 2\beta\sum_{i=2}^{n}[\ln X_i - \ln X_{i-1}] = 2\beta[\ln X_n - \ln X_1]$$

has a chi-squared distribution with $2(n-1)$ degrees of freedom. Furthermore, κ and ε are independent.

Meanwhile, denote

$$\xi = \frac{\varepsilon}{(n-1)\kappa} = \frac{\ln X_n - \ln X_1}{(n-1)(\ln X_1 - \ln\theta)},$$

and

$$\eta = \kappa + \varepsilon = 2\beta[\ln X_n - \ln\theta].$$

Note that ξ has an F distribution with $2(n-1)$ and 2 degrees of freedom, η has a chi-squared distribution with $2n$ degrees of freedom, and that ξ and η are independent (see Johnson et al. [10], p. 350).

Theorem 1. *Suppose that $X_1, X_2, \ldots, X_n$, are record values from two-parameter Pareto distribution (2), then for any $0 < \alpha < 1$, the $100(1-\alpha)\%$ confidence intervals for β and θ are given by*

$$\left(\frac{\chi^2_{2(n-1)}(1-\frac{\alpha}{2})}{2[\ln X_n - \ln X_1]}, \frac{\chi^2_{2(n-1)}(\frac{\alpha}{2})}{2[\ln X_n - \ln X_1]}\right),$$

and

$$\left(\exp\left\{\ln X_1 - \frac{\ln X_n - \ln X_1}{(n-1)F_{(2(n-1),2)}(1-\frac{\alpha}{2})}\right\}, \exp\left\{\ln X_1 - \frac{\ln X_n - \ln X_1}{(n-1)F_{(2(n-1),2)}(\frac{\alpha}{2})}\right\}\right),$$

where $F_{(2(n-1),2)}(p)$ is the $100p\%$ right-tail percentile of F distribution with $2(n-1)$ and 2 degrees of freedom.

Proof. Since ε and ξ have a chi-squared distribution and an F distribution, respectively, then

$$P\left(\chi^2_{2(n-1)}(1-\frac{\alpha}{2}) < \varepsilon < \chi^2_{2(n-1)}(\frac{\alpha}{2})\right) = 1-\alpha,$$

and

$$P\left(F_{(2(n-1),2)}(1-\frac{\alpha}{2}) < \xi < F_{(2(n-1),2)}(\frac{\alpha}{2})\right) = 1-\alpha.$$

The results can be yielded directly. □

Remark. In practical, We, sometimes, need to compare if the value of parameter is in accord with that of our past experience, this requires us to take some testing work. Using the properties of statistics ε and ξ, we can give some hypotheses testing here. Since function

$$\varepsilon(\beta; n) = 2\beta[\ln X_n - \ln X_1]$$

is strictly increasing in β. Thus to test the hypotheses $H_0 : \beta = \beta_0$ versus $H_\alpha : \beta > \beta_0$ (or $\beta < \beta_0$), the decision rule is to reject H_0 if $\varepsilon(\beta_0; n) > \chi^2_{2(n-1)}(\alpha)$ (or $\varepsilon(\beta_0; n) < \chi^2_{2(n-1)}(1-\alpha)$), here $\alpha \in (0, 1)$ is the testing level. Similarly, the two-side test $H_0 : \beta = \beta_0$ versus $H_\alpha : \beta \neq \beta_0$, the decision rule is to reject H_0 if $\varepsilon(\beta_0; n) > \chi^2_{2(n-1)}(\frac{\alpha}{2})$ (or $\varepsilon(\beta_0; n) < \chi^2_{2(n-1)}(1-\frac{\alpha}{2})$).

For the parameter θ, since function

$$\xi(\theta; n) = \frac{\ln X_n - \ln X_1}{(n-1)(\ln X_1 - \ln \theta)}$$

is strictly increasing in θ. For testing the hypotheses $H_0 : \theta = \theta_0$ versus $H_\alpha : \theta > \theta_0$ (or $\theta < \theta_0$), the decision rule is to reject H_0 if $\xi(\theta_0; n) > F_{(2(n-1),2)}(\alpha)$ (or $\xi(\theta_0; n) < F_{(2(n-1),2)}(1-\alpha)$). For two-side test $H_0 : \theta = \theta_0$ versus $H_\alpha : \theta \neq \theta_0$, the decision rule is to reject H_0 if $\xi(\theta_0; n) > F_{(2(n-1),2)}(\frac{\alpha}{2})$ (or $\xi(\theta_0; n) < F_{(2(n-1),2)}(1-\frac{\alpha}{2})$).

Theorem 2. *Suppose that $X_1, X_2, \ldots, X_n$, are record values from two-parameter Pareto distribution (2), then for any $0 < \alpha < 1$, a $100(1-\alpha)\%$ confidence region for (β, θ) is determined by following inequalities:*

$$\begin{cases} \exp\left(\ln X_1 - \frac{\ln X_n - \ln X_1}{(n-1)F_{(2(n-1),2)}(\frac{1+\sqrt{1-\alpha}}{2})}\right) < \theta < \exp\left(\ln X_1 - \frac{\ln X_n - \ln X_1}{(n-1)F_{(2(n-1),2)}(\frac{1-\sqrt{1-\alpha}}{2})}\right), \\ \left[\frac{\chi^2_{2n}(\frac{1+\sqrt{1-\alpha}}{2})}{2(\ln X_n - \ln \theta)}\right] < \beta < \left[\frac{\chi^2_{2n}(\frac{1-\sqrt{1-\alpha}}{2})}{2(\ln X_n - \ln \theta)}\right]. \end{cases}$$

Proof. Since ξ and η have an F distribution and a chi-squared distribution, respectively, we have

$$P\left(F_{(2(n-1),2)}(\frac{1+\sqrt{1-\alpha}}{2}) < \xi < F_{(2(n-1),2)}(\frac{1-\sqrt{1-\alpha}}{2})\right) = \sqrt{1-\alpha},$$

and

$$P\left(\chi^2_{2n}(\frac{1+\sqrt{1-\alpha}}{2}) < \eta < \chi^2_{2n}(\frac{1-\sqrt{1-\alpha}}{2})\right) = \sqrt{1-\alpha}.$$

From the independent property of ξ and η, the result can be obtained by using usual transformation techniques. □

3. Bayes estimation under symmetric and asymmetric loss

This section is devoted to investigate Bayes estimation for the parameters β and θ, as well as survival function $R(t)$ and hazard function $H(t)$, of the two-parameter Pareto distribution (2) under symmetric and asymmetric losses under record values.

3.1. Prior information and preliminary. In recent decades, the Bayes viewpoint, as a powerful and valid alternative to traditional statistical perspectives, has received frequent attention for statistical inference. In this paper we adapt a different method to derive the Bayes estimation under different loss functions. To be specific, here we suppose that θ has a discrete prior and β has a continuous conditional prior for given θ. That is to say, parameter θ follows a discrete prior distribution

$$P(\theta = \theta_j) = \eta_j, j = 1, 2, \ldots, N, \tag{13}$$

where $\eta_j, j = 1, 2, \ldots, N$, are positive numbers satisfying $\sum_{j=1}^{N} \eta_j = 1$.

For given θ, parameter β has a conditional conjunction prior distribution as follows

$$\pi(\beta|\theta_j) = a_j e^{-a_j\beta}, a_j > 0, \beta > 0, \tag{14}$$

where $a_j, j = 1, 2, \ldots, N$, are hyperparameters.

From (4) and (14), the conditional posterior distribution of β for given θ can be expressed as

$$\begin{aligned} \pi^*(\beta|x, \theta_j) &= \frac{\pi(\beta|\theta_j) f_X(x; \beta, \theta_j)}{\int_0^\infty \pi(\beta|\theta_j) f_X(x; \beta, \theta_j) d\beta} \\ &= \frac{(a_j + A_j)^{n+1}}{\Gamma(n+1)} \beta^n \exp\{-\beta(a_j + A_j)\}, \end{aligned} \tag{15}$$

where $A_j = \ln x_n - \ln \theta_j, j = 1, 2, \ldots, N$.

Meanwhile, from (4), (13) and (14), the joint posterior pdf of (β, θ) is given by

$$\begin{aligned} \pi^*(\beta, \theta_j|x) &= \frac{P(\theta = \theta_j)\pi(\beta|\theta_j) f(x|\beta, \theta_j)}{\sum_{j=1}^{N} \int_0^\infty P(\theta = \theta_j)\pi(\beta|\theta_j) f(x|\beta, \theta_j) d\beta} \\ &= \frac{\eta_j a_j \beta^n e^{-\beta(a_j + A_j)}}{\sum_{j=1}^{N} \eta_j a_j \Gamma(n+1)/[a_j + A_j]^{n+1}}. \end{aligned}$$

Furthermore, the marginal posterior distribution of θ_j can be written as

$$\begin{aligned} P_j &= P(\theta = \theta_j|x) = \int_0^\infty \pi^*(\beta, \theta_j|x) d\beta \\ &= \frac{\eta_j a_j [a_j + A_j]^{-(n+1)}}{\sum_{j=1}^{N} \eta_j a_j [a_j + A_j]^{-(n+1)}}. \end{aligned} \tag{16}$$

3.2. Bayes estimation under symmetric loss. In this subsection we consider the Bayes estimation for the reliability performances under square error loss function, which is one of very popular symmetric loss. The choice of squared error loss as loss function has twofold advantage of easy of computation and of leading to estimators that can be obtained directly. Since the Bayes estimation under squared error loss is to be the posterior expectation, using (15) and (16), the Bayes estimators for β, θ, as well as $R(t)$ and $H(t)$ at mission time t, namely $\hat{\beta}_{BS}, \hat{\theta}_{BS}, \hat{R}_{BS}(t)$, and $\hat{H}_{BS}(t)$, are given by

$$\begin{array}{ll}\hat{\beta}_{BS} & = \int_0^\infty \sum_{j=1}^N \beta\ P_j \pi^*(\beta|\theta_j, x) d\beta = \sum_{j=1}^N \frac{P_j(n+1)}{a_j+A_j},\\ \hat{\theta}_{BS} & = \sum_{j=1}^N P_j \theta_j,\\ \hat{R}_{BS} & = \int_0^\infty \sum_{j=1}^N R(t) P_j \pi^*(\beta|\theta_j, x) d\beta = \sum_{j=1}^N \frac{P_j(a+A_j)^{n+1}}{(a_j+A_j+\ln t-\ln\theta_j)^{n+1}},\\ \hat{H}_{BS} & = \int_0^\infty \sum_{j=1}^N H(t) P_j \pi^*(\beta|\theta_j, x) d\beta = \sum_{j=1}^N \frac{P_j(n+1)}{t(a_j+A_j)}.\end{array} \tag{17}$$

3.3. Bayes estimation under asymmetric loss. In statistical decision theory and Bayes analysis, the squared error loss function, as one of very popular symmetric loss, is widely used due to its great analysis property such as easy calculation. Under squared error loss, it is to be thought that squared error loss penalizes same importance to overestimation and underestimation. In some practical estimation and prediction problems, however, the overestimation and underestimation have different estimated risks. Thus using symmetric loss function may be inappropriate, and an asymmetric Linear-exponential (Linex) loss function has been introduced by Varian [20], and further illustrated by Zellner [22]. Several studies have been discussed on the use of the Linex loss. For instance, Akdeniz [2] generalized Liu estimator and obtained a new biased estimator when Linex loss is used. Hoque et al. [9] studied the performance of the unrestricted estimator and preliminary test estimator of the slope parameter of simple Linear regression model under Linex loss. The Linex loss function is in the performance of approximately linearly on one side of zero and approximately exponentially on the other side, which is defined as follows

$$L(\delta, \phi(\theta)) = e^{c(\delta-\phi(\theta))} - c(\delta - \phi(\theta)) - 1, c \neq 0, \tag{18}$$

where δ is an estimator of $\phi(\theta)$.

From (18), it is seen that, when $c > 0$, overestimation is more serious than underestimation; When $c < 0$, the conclusion is opposite. As c nears to zero, the Linex loss function is approximately the squared error loss, and therefore almost symmetric. A review of Linex loss and its properties are investigated by Parsian and Kirmani [15]. The Bayes estimator under Linex loss is given by

$$\hat{\phi}_{BL} = -\frac{1}{c} \ln[E_\phi(e^{-c\phi(\theta)})], \tag{19}$$

provided that the expectation $E_\phi(e^{-c\phi(\theta)})$ exists and is finite, where $E_\phi(\cdot)$ denotes posterior expectation with respect to the posterior density of $\phi(\theta)$.

Using (15), (16) and (19), the Bayes estimators for β, θ, as well as $R(t)$ and $H(t)$ at mission time t, namely $\hat{\beta}_{BL}, \hat{\theta}_{BL}, \hat{R}_{BL}(t)$, and $\hat{H}_{BS}(t)$, under Linex loss can be written as

$$\begin{aligned} \hat{\beta}_{BL} &= -\frac{1}{c}\ln\left[\int_0^\infty \sum_{j=1}^{n} P_j e^{-c\beta}\pi^*(\beta|\theta_j,x)d\beta\right] \\ &= -\frac{1}{c}\ln\left[\sum_{j=1}^{N} \frac{P_j(a_j+A_j)^{n+1}}{(a_j+c+A_j)^{n+1}}\right], \\ \hat{\theta}_{BL} &= -\frac{1}{c}\ln\left[\sum_{j=1}^{N} P_j e^{-c\theta_j}\right], \\ \hat{R}_{BL}(t) &= -\frac{1}{c}\ln\left[\sum_{j=1}^{N}\sum_{k=0}^{\infty} \frac{P_j}{k!}\frac{[-c]^k(a_j+A_j)^{n+1}}{(a_j+A_j+k\ln t-k\ln\theta_j)^{n+1}}\right], \\ \hat{H}_{BL}(t) &= -\frac{1}{c}\ln\left[\sum_{j=1}^{N} \frac{P_j(a_j+A_j)^{n+1}}{(a_j+A_j+c/t)^{n+1}}\right]. \end{aligned} \tag{20}$$

Another useful asymmetric loss function is the General Entropy (GE) loss, which is defined as follows

$$L(\delta,\phi(\theta)) \propto \left(\frac{\delta}{\phi(\theta)}\right)^q - q\ln\left(\frac{\delta}{\phi(\theta)}\right) - 1.$$

When $q > 0$, the positive error ($\delta > \phi(\theta)$) causes more serious consequences than that caused by a negative error, and vice versa. When $q = 1$, the GE reduces to the conventional entropy loss function. The Bayes estimator under GE loss is given by

$$\hat{\phi}_{BG} = \left(E_\phi[\phi(\theta)^{-q}]\right)^{-1/q}, \tag{21}$$

provided that $E_\phi[\phi(\theta)^{-q}]$ exists and is finite.

Using (15), (16) and (21), the Bayes estimators for β,θ, as well as $R(t)$ and $H(t)$ at mission time t, namely $\hat{\beta}_{BG},\hat{\theta}_{BG},\hat{R}_{BG}(t)$, and $\hat{H}_{BG}(t)$, under GE loss can be expressed as

$$\begin{aligned} \hat{\beta}_{BG} &= \left[\int_0^\infty \sum_{j=1}^{N} P_j\beta^{-q}\pi^*(\beta|\theta_j,x)d\beta\right]^{-1/q} \\ &= \left[\sum_{j=1}^{N} \frac{\Gamma(n+1-q)}{\Gamma(n+1)} P_j(a_j+A_j)^q\right]^{-1/q}, \\ \hat{\theta}_{BG} &= \left[\sum_{j=1}^{N} P_j\theta_j^{-q}\right]^{-1/q}, \\ \hat{R}_{BG}(t) &= \left[\sum_{j=1}^{N} \frac{P_j(a_j+A_j)^{n+1}}{(a_j+A_j+q\ln\theta_j-q\ln t)^{n+1}}\right]^{-1/q}, \\ \hat{H}_{BG}(t) &= \left[\sum_{j=1}^{N} \frac{\Gamma(n+1-q)}{\Gamma(n+1)} P_j t^q(a_j+A_j)^q\right]^{-1/q}. \end{aligned} \tag{22}$$

where $n+1-q > 0$ for the existence of Bayes estimators.

3.4. The choice of hyperparameters. Sometimes it is not always possible to know the exact value of the hyper-parameters a_j in prior, the estimation problem is considered for the unknown hyperparameters $a_j, j = 1,2,\ldots,N$ in this subsection

In the classic maximum likelihood estimation method, the MLEs of β,θ are the values for which the likelihood function is largest over all possibilities. This method is used for the maximum points of density. In the view of the Bayes approach, it is an effective measure to estimate the unknown parameter by making use of expectation and MLE, which could utilize the prior information properly.

Here we adopt a similar way to derive the estimators of the hyperparameters as Soliman [17, 18] by using expectation and MLE.

Recall (6) and (7), the MLEs of $R(t)$ and $H(t)$ at mission time t are given by

$$\hat{R}(t) = \hat{\theta}_M^{\hat{\beta}_M} t^{-\hat{\beta}_M}, \quad \hat{H}(t) = \hat{\beta}_M / t.$$

Meanwhile, the expectation of $R(t)$ can be expressed as

$$ER(t) = \int_0^\infty R(t)\pi(\beta|\theta_j)d\beta = \frac{a_j}{a_j + \ln t - \ln\theta_j}.$$

For a given mission time t, let $ER(t) = \hat{R}(t)$, the estimators of a_j, namely $\hat{a}_j$, can be written as

$$\hat{a}_j = \frac{\hat{R}(t)}{1 - \hat{R}(t)}\left[\ln t - \ln\theta_j\right], j = 1, 2, \ldots, N.$$

Another useful alternative method to estimate the hyperparameters $a_j, j = 1, 2, \ldots, N$, is the maximum likelihood type II estimation, or simply ML-II method (see Berger, [6], p. 99).

Let $T_i = \ln X_i - \ln\theta$, then $T_i, i = 1, 2, \ldots, n$, are the upper record values from the exponential distribution with conditional density function $f_T(t;\beta) = \beta\exp\{-\beta t\}, t > 0$. Furthermore, for given θ_j, the marginal density function, as well as cdf of T_i, are give by

$$f_T(t) = \int_0^\infty \pi(\beta|\theta_j)f_T(t;\beta)d\beta = \frac{a_j}{(t + a_j)^2}, t > 0,$$

and

$$F_T(t) = \int_0^t f_T(x)dx = 1 - \frac{a_j}{t + a_j}, t > 0.$$

From (1), the joint pdf of $(T_1, T_2, \ldots, T_n)$ can be expressed as

$$\begin{aligned} f(t; a_j) &= f_T(t_n)\prod_{i=1}^{n-1}\frac{f_T(t_i)}{1 - F_T(t_i)} = \frac{a_j}{t_n + a_j}\prod_{i=1}^{n}\frac{1}{t_i + a_j} \\ &= \frac{a_j}{a_j + \ln x_n - \ln\theta_j}\prod_{i=1}^{n}\frac{1}{a_j + \ln x_i - \ln\theta_j}. \end{aligned}$$

Furthermore, the log-likelihood function can be written as

$$\begin{aligned} L(a_j; x) &= \ln f(t; a_j) \\ &= \ln a_j - \ln(a_j + \ln x_n - \ln\theta_j) - \sum_{i=1}^{n}\ln(a_j + \ln x_i - \ln\theta_j), \end{aligned}$$

taking derivative for $L(a_j; x)$ with respect to a_j, then

$$\frac{\partial}{\partial a_j}L(a_j; x) = \frac{1}{a_j} - \frac{1}{a_j + \ln x_n - \ln\theta_j} - \sum_{i=1}^{n}\frac{1}{a_j + \ln x_i - \ln\theta_j}.$$

As the ML-II estimator of a_j is needed, we just need to draw a conclusion that the equation $\frac{\partial}{\partial a_j}L(a_j;x)=0$ has only one root with respect to a_j for all $j=1,2,\ldots,N$.

Denote

$$g_1(a_j)=\frac{1}{a_j}-\frac{1}{a_j+\ln x_n-\ln\theta_j},\quad g_2(a_j)=\sum_{i=1}^{n}\frac{1}{a_j+\ln x_i-\ln\theta_j}.$$

Since

$$\begin{aligned}\lim_{a_j\to 0}g_1(a_j) &= +\infty,\quad \lim_{a_j\to+\infty}g_1(a_j)=0,\\ \frac{\partial}{\partial a_j}g_1(a_j) &= -\frac{1}{a_j^2}+\frac{1}{(a_j+\ln x_n-\ln\theta_j)^2}<0,\\ \frac{\partial^2}{\partial a_j^2}g_1(a_j) &= \frac{2}{a_j^3}-\frac{2}{(a_j+\ln x_n-\ln\theta_j)^3}>0,\end{aligned}$$

and

$$\begin{aligned}\lim_{a_j\to 0}g_2(a_j) &= \sum_{i=1}^{n}\frac{1}{\ln x_i-\ln\theta_j},\quad \lim_{a_j\to+\infty}g_2(a_j)=0,\\ \frac{\partial}{\partial a_j}g_2(a_j) &= -\sum_{i=1}^{n}\frac{1}{(a_j+\ln x_i-\ln\theta_j)^2}<0,\\ \frac{\partial^2}{\partial a_j^2}g_2(a_j) &= \sum_{i=1}^{n}\frac{2}{(a_j+\ln x_i-\ln\theta_j)^3}>0.\end{aligned}$$

It is noted that both functions of $g_1(a_j)$ and $g_2(a_j)$ are strictly monotone decreasing concave functions.

Furthermore, since

$$\begin{aligned}\lim_{a_j\to\infty}\frac{g_2(a_j)}{g_1(a_j)} &= \sum_{i=1}^{n}\lim_{a_j\to\infty}\frac{1}{a_j+\ln x_i-\ln\theta_j}\,\frac{a_j(a_j+\ln x_n-\ln\theta_j)}{\ln x_n-\ln\theta_j}\\ &= \sum_{i=1}^{n}\lim_{a_j\to\infty}\frac{2a_j+\ln x_n-\ln\theta_j}{\ln x_n-\ln\theta_j}=\infty,\end{aligned}$$

then the equation $\frac{\partial}{\partial a_j}L(a_j;x)=0$ has only one root, which implies the ML-II estimator of a_j is unique. Since there is no closed solution for a_j, the estimator of a_j, say $\hat{a}_j$, can be derived using following iterative formula

$$\frac{1}{a_j^{(k+1)}}=\frac{1}{a_j^{(k)}+\ln x_n-\ln\theta_j}+\sum_{i=1}^{n}\frac{1}{a_j^{(k)}+\ln x_i-\ln\theta_j},k=0,1,2,\cdots$$

for all $j=1,2,\ldots,N$, where $a_j^{(k)}$ is the kth iterative value, and $a_j^{(0)}$ is an initial value.

4. Numerical examples

In this section, two examples are given to illustrate the results provided in previous sections. We apply the proposed methods to one of practical data set and another simulated data set. Further, a Monte Carlo simulation is conducted to compare the simulation results.

Example 1. (Real-life data) The following upper record values which represent the values of the average July temperatures (in degrees centigrade) of Neuenburg, Switzerland, during the period 1864-1993 (from Kluppelberg and Schwere [11]).

19.0 20.1 21.4 21.7 22.0 22.1 22.6 23.4

Arnold and Press [5] have showed that the Pareto distribution is a reasonable model for this data set.

From (6) and (7), the MLEs for the parameters θ and β are given by $\hat{\theta}_M = 19.0$ and $\hat{\beta}_M = 38.4067$, respectively.

Using following percentiles

$$\chi^2_{14}(0.95) = 6.5706, \quad \chi^2_{14}(0.05) = 23.6848,$$
$$F_{(14,2)}(0.95) = 0.2675, \quad F_{(14,2)}(0.05) = 19.4244.$$

By Theorem 1, the 90% confidence intervals for β and θ are (15.7720,56.8526) and (16.9998,18.9709), respectively.

Furthermore, using the percentiles as follows

$$\chi^2_{16}(0.9743) = 6.9454, \quad \chi^2_{16}(0.0257) = 28.7469,$$
$$F_{(14,2)}(0.9743) = 0.2079, \quad F_{(14,2)}(0.0257) = 38.3370.$$

By Theorem 2, the 90% confidence region for (β, θ) is determined from the following inequalities:

$$\begin{cases} 16.4662 < \theta < 18.9853, \\ \left[\frac{3.4727}{3.1527 - \ln\theta}\right] < \beta < \left[\frac{19.1685}{3.1527 - \ln\theta}\right]. \end{cases}$$

From above description of the confidence region, one can clearly find that the confidence region is large when θ is large.

Example 2. (Simulated data) In order to give a simulation study, here we provide an algorithm to generate a group of record values as following steps:

Step 1. Generating a group of i.i.d. samples, namely $Z_1, Z_2, \ldots, Z_n$, from uniform distribution with density function $f(z) = 1, 0 < z < 1$ and $f(z) = 0$ otherwise.

Step 2. Making transformation $Y_i = -\ln(1 - Z_i)$, then $Y_i, i = 1, 2, \ldots, n$, are the i.i.d. samples from standard exponential distribution $Exp(1)$ with density function $f(y) = 1 - e^{-y}, 0 < y < \infty$.

Step 3. Let $W_i = Y_1 + Y_2 + \ldots + Y_i$, since the exponential distribution has he lack of memory property, the sequences $W_i, i = 1, 2, \ldots, n$, are the record values from standard exponential distribution.

Step 4. Denote $U_i = 1 - e^{-W_i}$, then sequence $U_i, i = 1, 2, \ldots, n$, are the record values from uniform distribution with density $f(u) = 1, 0 < u < 1$ and $f(u) = 0$ otherwise.

Step 5. For arbitrary cdf $F(x)$, making transformation $X_i = F^{-1}(U_i)$, sequence $X_i, i = 1, 2, \ldots, n$, are record values sequence with cdf $F(x)$, where $F^{-1}(\cdot)$ is the inverse function of $F(\cdot)$.

Using the algorithm described above, for given $\beta = 0.9, \theta = 1.6, n = 5$, the following record values data of the two-parameter Pareto distribution are simulated:

2.435 6.850 7.965 8.086 24.583

Using following percentiles

$$\chi^2_8(0.95) = 2.7326, \quad \chi^2_8(0.05) = 15.5073,$$
$$F_{(8,2)}(0.95) = 0.2243, \quad F_{(8,2)}(0.05) = 19.3710.$$

By Theorem 1, the 90% confidence intervals for β and θ are (0.5912,3.3530) and (0.1844,2.3634), respectively.

Meanwhile, using the percentiles as follows

$$\chi^2_{10}(0.9743) = 3.2713, \quad \chi^2_{10}(0.0257) = 20.3986,$$
$$F_{(8,2)}(0.9743) = 0.1669, \quad F_{(8,2)}(0.0257) = 38.2835.$$

By Theorem 2, the 90% confidence region for (β, θ) is determined from the following inequalities:

$$\begin{cases} 0.0764 < \theta < 2.3985, \\ \left[\frac{1.6356}{3.2021 - \ln\theta}\right] < \beta < \left[\frac{10.1993}{3.2021 - \ln\theta}\right]. \end{cases}$$

Fig. 1 shows the 90% confidence region for (β, θ). It is clear to find that the region is large when θ is large.

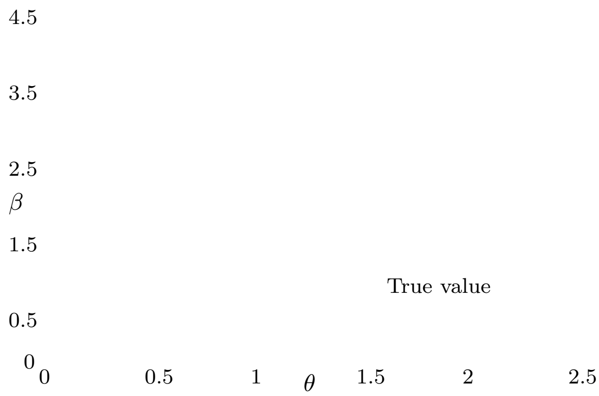

Fig. 1. A 90% confidence region for (β, θ).

In order to examine how well the proposed approach works for constructing confidence intervals and regions, we will make 5000 times simulation for the estimation of the exact and approximate (Appro) confidence intervals and regions in terms of convergence probabilities. The simulation results are summarized in Table 1.

TABLE 1. Coverage probabilities of interval and region estimates when $(\beta,\theta)=(0.9,1.6)$.

n	β		θ		(β,θ)	
	Exact	Appro	Exact	Appro	Exact	Appro
3	0.865	0.797	0.861	0.799	0.919	0.751
5	0.903	0.856	0.905	0.873	0.914	0.899
7	0.896	0.820	0.910	0.865	0.926	0.857
9	0.914	0.868	0.907	0.869	0.918	0.873

Using the simulated data and given θ_j,η_j with $N=10, t=3$, Table 2 summarized the values of $\hat{a}_j, A_j, P_j, j=1,2,\ldots,N$. The Bayes estimates, as well as MLEs, for the parameter β,θ, survival function $R(t)$ and hazard function $H(t)$ are computed from (17), (20) and (22), and the calculated results are listed in Table 3.

TABLE 2. Prior information and posterior probabilities $(n=5, t=3)$

j	1	2	3	4	5	6	7	8	9	10
θ_j	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
η_j	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
$\hat{a}_j$	1.003	0.916	0.836	0.762	0.693	0.628	0.568	0.511	0.457	0.405
A_j	3.106	3.020	2.940	2.865	2.796	2.732	2.671	2.614	2.560	2.509
P_j	0.049	0.058	0.067	0.079	0.091	0.103	0.117	0.130	0.144	0.157

Note: here the hyperparameters are estimated by using the first method.

TABLE 3. Bayes estimates for reliability index (n=5, t=3)

Loss and index		$\hat{\beta}$	$\hat{\theta}$	$\hat{R}(t)$	$\hat{H}(t)$
MLE		0.872	2.434	0.833	0.291
Squared loss		1.775	1.580	0.360	0.592
Linex loss	c=2	1.572	2.043	0.351	0.549
	c=-1	1.622	1.719	0.638	0.626
GE loss	q=1	1.752	1.520	0.732	0.387
	q=-2	1.371	1.607	0.469	0.358

In order to illustrate the accuracy of the different estimators, repeating 1000 simulations as above, the estimated risk (ER) are computed as the average of their squared deviations. The expression is

$$\frac{1}{1000}\sum_{i=1}^{1000}(\hat{\phi}-\phi)^2,$$

where ϕ and $\hat{\phi}$ denote the original value and Bayes estimates of $\beta, \theta, R(t)$ and $H(t)$, respectively. The estimated risk of Bayes estimates under different loss are listed in Table 4.

TABLE 4. The estimated risk of Bayes estimates under different loss(t=3,c=2,q=1)

MLE / Sample size	β	θ	$R(t)$	$H(t)$
n=3	0.675	0.924	0.575	0.583
n=5	0.383	0.673	0.351	0.321
n=7	0.272	0.218	0.124	0.127
Squared Loss / Sample size	β	θ	$R(t)$	$H(t)$
n=3	0.975	0.515	0.657	0.521
n=5	0.813	0.248	0.339	0.275
n=7	0.481	0.096	0.103	0.096
Linex Loss / Sample size	β	θ	$R(t)$	$H(t)$
n=3	0.576	0.617	0.528	0.496
n=5	0.457	0.423	0.201	0.219
n=7	0.101	0.099	0.084	0.085
GE Loss / Sample size	β	θ	$R(t)$	$H(t)$
n=3	0.876	0.478	0.343	0.618
n=5	0.699	0.123	0.208	0.397
n=7	0.211	0.078	0.101	0.142

5. Conclusions

In this paper the estimation problem are considered for the parameters as well as the reliability and hazard functions of the Pareto model based on records by using Bayes and non-Bayes procedures. There are some conclusions which have been noticed as follows

1: Excepting for being used to analysis social and economical phenomenon, the two-parameter Pareto distribution has received more and more attentions for its application in reliability theory, quality control duration and failure time modeling, as well as other related fields.

2: The record values is thought of as a special topic in order statistics, which has been developed widely in applications such as reliability theory, meteorology, sports analysis, hydrology, and stock market analysis.

3: From Table 1, it is noted that, under different sample size, the coverage probabilities of the exact confidence intervals for β and θ, as well as the exact confidence region for (β, θ) are all close to the desired level of 0.90. Comparing with the coverage probabilities of the approximate estimates, the exact confidence intervals and regions are better than the asymptotic ones.

4: From Table 3 and Table 4, It will be observed that the Bayes estimates are superior to MLEs for the parameters β, θ, as well as survival function $R(t)$ and hazard function $H(t)$, respectively. The results show that the asymmetric Bayes estimates (under Linex and GE losses) are general better than MLEs, which makes them more attractive for using in practical problem. Meanwhile, the estimated risks of the Bayes estimators get smaller under both symmetric and asymmetric losses with increasing sample size.

5: It may be remarked that there are two methods mentioned in this paper to estimate the hyperparameters. Since the first one depends on the value of MLEs of the parameters, which, however, sometimes cannot provide a good enough estimates, hence the authors tend to recommend the second method, in which the discrete prior information are also utilized there.

References

1. M. Ahsanullah, *Introduction to record statistics*, New York, NOVA Science, Huntington, 1995.
2. F. Akdeniz, *New biased estimators under the Linex loss function*, Statistical Papers *45*(2004), 175-190.
3. E.K. Al-Hussaini, A.A. Ahmed, *On Bayesian interval prediction of future records*, Test *12*(2003), 79-99.
4. B.C. Arnold, N. Balakrishnan, H.N., Nagaraja, *Records*, New York, John Wiley & Sons, 1998.
5. B.C. Arnold, S.J. Press, *Bayesian estimation and prediction for pareto data*, Journal of the American Statistical Association *84*(1989), 1079-1084.
6. J.O. Berger, *Statistical Decision Theory and Bayesian Analysis (2nd edition)*, New York, Springer, 1985.
7. K.N. Chandler, *The distribution and frequency of record values*, Journal of the Royal Statistical Society, Series B *14*(1952), 220-228.
8. P.S. Dayna, V.H. Humberto, B.C., Arnold, *A goodness of fit test for the Pareto distribution in the presence of Type II censoring based on the cumulative hazard function*, Computational Statistics & Data Analysis *54*(2010), 833-842.
9. Z. Hoque, S. Khan, J. Wesolowski, *Performance of preliminary test estimator under linex loss function*, Communications in Stattistics: Theory and Methods *38*(2)(2009), 252-261.
10. N.L. Johnson, S. Kota, N. Balakrishnan, *Continuous univariate distributions, Vol1, 2nd ed.* New York, John Wiley & Sons, 1994.

11. C. Kluppelberg, P. Schwere, *Records in time series: an investigation to global warming, Berichte Zur Stachastik und Verwandten Gebieten*, Johannes Gutenberg-Universitt, Mainz, Tech. Rep. 95-4, June 1995.
12. T.M. Mohamed, Z.R. Mohamed, *Bayesian prediction of temperature records using the Pareto model*, Environmetrics *15*(2004), 701-710.
13. V.B. Nevzorov, *Record: Mathematical Theory.* Translation of Mathematical Monographs, American Mathematical Society, vol. 194, Providence, RI (2001).
14. V. Pareto, *Cours d'Economie Politique*, Rouge et Cie, Paris (1897).
15. A. Parsian, S.N.U.A. Kirmani, *Estimation under Linex loss function.* In Handbook of Applied Econometrics and Statistical Inference. Eds. Aman Ullah, Alan T. K. Wan, and Anoop Chaturvedi. Marcel Dekker, Inc. (2002), 53-76
16. A.A. Soliman, *Bayes prediction in a pareto lifetime models with random sample size*, Journal of the Royal Statistical Society, Series D *49*(1)(2000), 51-62.
17. A.A. Soliman, *Estimation of parameters of life from progressively censored data using Burr-XII Model.* IEEE Transactions on Reliability *54*(1)(2005), 34-42.
18. A.A. Soliman, *Estimations for pareto model using general progressive censored data and asymmetric loss*, Communications in Stattistics: Theory and Methods *37*(2008), 1353-1370.
19. R.C. Tiwari, Y. Yang, J.N. Zalkikar, *Bayes estimation for Pareto failure model using Gibbs sampling*, IEEE Transactions on Reliability *45*(1996), 471-476.
20. H.R. Varian, *A Bayesian approach to real estate assessment*, In Studies Bayesian Econometrics and Statistics in Honor of Leonard J. Savage. Fienberg, S.E. and Zellner, A., Eds. North Holland: Amsterdam (1975), 195-208.
21. J.W. Wu, W.C. Lee, S.C. Chen, *Computational comparison of prediction future lifetime of electronic components with Pareto distribution based on multiply type II censored samples* Applied Mathematics and Computation, *184*(2007), 374-406.
22. A. Zellner, *Bayesian estimation and prediction using asymmetric loss functions*, Journal of the American Statistical Association *81*(1986), 446-451.

Liang Wang is a Ph. D. student in Department of Applied Mathematics, Northwestern Polytechnical University, Xian, China. He received M.Sc. from Northwestern Polytechnical University in 2008. His research interests focus on applied probability and statistics, Bayes theory and reliability analysis.

Department of Applied Mathematics, Northwestern Polytechnical University, Xian, Shaanxi 710072, P.R. China.
e-mail: liang610112@gmail.com

Yimin Shi is a professor at Department of Applied Mathematics, Northwestern Polytechnical University, Xian, China. His research interests are applied probability and statistics and reliability analysis.

Department of Applied Mathematics, Northwestern Polytechnical University, Xian, Shaanxi 710072, P.R. China.
e-mail: ymshi@nwpu.edu.cn

Ping Chang is a M.Sc. student from Xian Jiaotong University. Her research interests are clinic statistics and Gynecologic Oncology.

College of Medicine, Xian Jiaotong University, Xi'an, Shaanxi 710061, P.R. China.
e-mail: ping610113@163.com