Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.12785/jsap/020310

Bayesian Inference and Prediction using Progressive First-Failure Censored from Generalized Pareto Distribution

Mohamed A. W. Mahmoud[1], *Ahmed A. Soliman*[2], *Ahmed H. Abd-Ellah*[3] *and Rashad M. El-Sagheer*[1,*]

[1]Mathematics Department, Faculty of Science, Al-Azhar University, Nasr city (11884), Cairo, Egypt
[2]Mathematics Department, Faculty of Science, Islamic University, Madinah, Saudi Arabia
[3]Mathematics Department, Faculty of Science, Sohag University, Sohag 82524, Egypt

Received: 7 Jun. 2012, Revised: 21 Sep. 2012, Accepted: 23 Sep. 2012
Published online: 1 Nov. 2013

Abstract: This paper describes the Bayesian inference and prediction of the generalized Pareto (GP) distribution for progressive first-failure censored data. We consider the Bayesian inference under a squared error loss function. We propose to apply Gibbs sampling procedure to draw Markov Chian Monte Carlo (MCMC) samples, and they have in turn, been used to compute the Bayes estimates with the help of importance sampling technique. We have performed a simulation study in order to compare the proposed Bayes estimators with the maximum likelihood estimators. We further consider two sample Bayes prediction to predicting future order statistics and upper record values from generalized Pareto (GP) distribution based on progressive first-failure censored data. The predictive densities are obtained and used to determine prediction intervals for unobserved order statistics and upper record values. A simulated data set is used to illustrate the results derived.

Keywords: Generalized Pareto distribution, Progressive first-failure censored sample, Bayesian estimations, Gibbs sampling, Markov Chain Monte Carlo, Posterior predictive density.

1 Introduction

Censoring is common in life-distribution work because of time limits and other restrictions on data collection. Censoring occurs when exact lifetimes are known only for a portion of the individuals or units under study, while for the remainder of the lifetimes information on them is partial. There are several types of censored tests. One of the most common censored test is type II censoring. It is noted that one can use type II censoring for saving time and money. However, when the lifetimes of products are very high, the experimental time of a type II censoring life test can be still too long. A generalization of type II censoring is progressive type II censoring, which is useful when the loss of live test units at points other than the termination point is unavoidable. Recently, the type II progressively censoring scheme has received considerable interest among the statisticians. See for example, Kundu [1] and Raqab et al. [2]. For the theory methods and applications of progressive censoring, one can refer to the monograph by Balakrishnan and Aggarwala [3] and the recent survey paper by Balakrishnan [4].

Johnson [5] described a life test in which the experimenter might decide to group the test units into several sets, each as an assembly of test units, and then run all the test units simultaneously until occurrence the first failure in each group. Such a censoring scheme is called a first-failure censoring scheme. Jun et al. [6] discussed a sampling plan for a bearing manufacturer. The bearing test engineer decided to save test time by testing 50 bearings in sets of 10 each. The first-failure times from each group were observed. Wu et al. [7] and Wu and Yu [8] obtained maximum likelihood estimates (MLEs), exact confidence intervals and exact confidence regions for the parameters of the Gompertz and Burr type XII distributions based on first-failure censored sampling, respectively. If an experimenter desires to remove some sets of test units before observing the first failures in these sets this life test plan is called a progressive first-failure censoring scheme which recently introduced by Wu and Kuş [9].

* Corresponding author e-mail: Rashadmath@Yahoo.com

In many practical problems of statistics, one wishes to use the results of previous data to predict a future observation from the same population. One way to do this is to construct an interval which will contain the future observation with a specified probability. This interval is called a prediction interval. Prediction has been applied in medicine, engineering, business, and other areas as well. Hahn and Meeker [10] have recently discussed the usefulness of constructing prediction intervals. Bayesian prediction bounds for future observations based on certain distributions have been discussed by several authors. Bayesian prediction bounds for observables having the Burr type XII distribution were obtained by Nigm [11], AL-Hussaini and Jaheen ([12],[13]), and Ali Mousa and Jaheen ([14],[15]). Burr type X distribution described by Jaheen and AL-Matrafi [16]. Lomax distribution described by Abd Ellah ([17],[18]).

Recently, Alamm et al. [19] obtained Bayesian prediction intervals for future order statistics from the generalized exponential distribution. Kundu and Howlader [20] studied Bayesian inference and prediction of inverse Weibull distribution for type II censored data. Ahmadi et al. [21] considered the Bayesian prediction of order statistics based on k-record values from exponential distribution. Ahmadi and MirMostafaee [22] obtained prediction intervals for order statistics as well as for the mean life time from a future sample based on observed usual records from an exponential distribution. Ali Mousa and Al-Sagheer [23] discussed the prediction problems for the Rayleigh based on progressively type II censored data.

Arandom variable X is said to have generalized Pareto (GP) distribution, if its probability density function (pdf) is given by

$$f_{(\zeta,\mu,\sigma)} = \frac{1}{\sigma}\left(1+\zeta\frac{x-\mu}{\sigma}\right)^{-(1/\zeta+1)}$$

where $\mu,\zeta \in \mathbb{R}$ and $\sigma \in (0,+\infty)$. For convenience, we reparametrized this distribution by defining $\sigma/\zeta = \beta, 1/\zeta = \alpha$ and $\mu = 0$. Therefore,

$$f(x) = \alpha\beta^{\alpha}(x+\beta)^{-(\alpha+1)}, \qquad x>0,\alpha,\beta>0. \tag{1}$$

The cumulative distribution function (cdf) is defined by

$$F(x) = 1-\beta^{\alpha}(x+\beta)^{-\alpha}, \qquad x>0,\alpha,\beta>0, \tag{2}$$

here α and β are the shape and scale parameters, respectively. It is also well known that this distribution has decreasing failure rate property. This distribution is also known as Pareto distribution of type II or Lomax distribution. This distribution has been shown to be useful for modeling and analizing the life time data in medical and biological sciences, engineering, etc. So, it has been received the greatest attention from theoretical and applied statisticians primarily due to its use in reliability and lifetesting studies. Many statistical methodes have been developed for this distribution, for a review of Pareto distribution of type II or Lomax distribution see Habibullh and Ahsanullah [24], Upadhyay and Peshwani [25] and Abd Ellah ([17],[18]) and references of them. Agreat deal of research has been done on estimating the parameters of a Lomax using both classical and Bayesian techniques.

In this paper first we consider the Bayesian inference of the shape and scale parameters for progressive first-failure censored data when both parameters are unknown. We assumed that the shape parameter α and the scale parameter β have the gamma prior and they are independently distributed. As expected in this case also, the Bayes estimates can not be obtained in closed form. We propose to use the Gibbs sampling procedure to generate MCMC samples, and then using the importance sampling methodology, we obtain the Bayes estimates of the unknown parameters. We perform some simulation experiments to see the behavior of the proposed Bayes estimators and compare their performances with the maximum likelihood estimators (MLEs).

Another important problem in life-testing experiments namely the prediction of unknown observables belonging to a future sample, based on the current available sample, known in the literature as the informative sample. For different application areas and for references, the readers are referred to AL-Hussaini [26]. In this paper we consider the prediction problem in terms of the estimation of the posterior predictive density of a future observation for two-sample prediction. We also construct predictive interval for a future observation using Gibbs sampling procedure. An illustrative example has been provided.

The rest of this paper is organized as follows: In Section 2, we describe the formulation of a progressive first-failure censoring scheme. In Section 3, we cover Bayes estimates of parameters using MCMC technique with the help of importance sampling technique. Monte Carlo simulation results are presented in Section 4. Bayes prediction for future order statistic and upper record values are provided in Section 5. and Section 6, respectively. Data analysis is provided in Section 7, and finally we conclude the paper in Section 8.

2 A progressive first-failure censoring scheme

In this section, first-failure censoring is combined with progressive censoring as in Wu and Kuş [9]. Suppose that n independent groups with k items within each group are put on a life test, R_1 groups and the group in which the first failure is observed are randomly removed from the test as soon as the first failure (say $X^{\mathbf{R}}_{1:m:n:k}$) has occurred, R_2 groups and the group in which the second first failure is observed are randomly removed from the test when the second failure (say $X^{\mathbf{R}}_{2:m:n:k}$) has occurred, and finally R_m $(m \leq n)$ groups and the group in which the $m-th$ first failure is observed are randomly removed from the test as soon as the $m-th$ failure (say $X^{\mathbf{R}}_{m:m:n:k}$) has occurred. The $X^{\mathbf{R}}_{1:m:n:k} < X^{\mathbf{R}}_{2:m:n:k} < \ldots < X^{\mathbf{R}}_{m:m:n:k}$ are called progressively first-failure censored order statistics with the progressive censoring scheme $\mathbf{R} = (R_1, R_2, \ldots, R_m)$. It is clear that m is number of the first-failure observed $(1 < m \leq n)$ and $n = m + R_1 + R_2 + \ldots + R_m$. If the failure times of the $n \times k$ items originally in the test are from a continuous population with distribution function $F(x)$ and probability density function $f(x)$, the joint probability density function for $X^{\mathbf{R}}_{1:m:n:k}, X^{\mathbf{R}}_{2:m:n:k}, \ldots, X^{\mathbf{R}}_{m:m:n:k}$ is given by

$$f_{1,2,\ldots,m}(x^{\mathbf{R}}_{1:m:n:k}, x^{\mathbf{R}}_{2:m:n:k}, \ldots, x^{\mathbf{R}}_{m:m:n:k}) = Ak^m \prod_{j=1}^{m} f(x^{\mathbf{R}}_{j:m:n:k})(1 - F(x^{\mathbf{R}}_{j:m:n:k}))^{k(\mathbf{R}_j+1)-1} \tag{3}$$

$$0 < x^{\mathbf{R}}_{1:m:n:k} < x^{\mathbf{R}}_{2:m:n:k} < \ldots < x^{\mathbf{R}}_{m:m:n:k} < \infty, \tag{4}$$

where

$$A = n(n - R_1 - 1)(n - R_1 - R_2 - 1)\ldots(n - R_1 - R_2 - \ldots R_{m-1} - m + 1). \tag{5}$$

Special cases

It is clear from (3) that the progressive first-failure censored scheme containing the following censoring schemes as special cases:

1. The first-failure censored scheme when $\mathbf{R} = (0, 0, \ldots, 0)$.
2. The progressive type II censored order statistics if $k = 1$.
3. Usually type II censored order statistics when $k = 1$ and $R = (0, 0, \ldots, n - m)$.
4. The complete sample case when $k = 1$ and $\mathbf{R} = (0, 0, \ldots, 0)$.

Also, It should be noted that $X^{\mathbf{R}}_{1;m,n,k}, X^{\mathbf{R}}_{2;m,n,k}, \ldots, X^{\mathbf{R}}_{m;m,n,k}$ can be viewed as a progressive type II censored sample from a population with distribution function $1 - (1 - F(x))^k$. For this reason, results for progressive type II censored order statistics can be extend to progressive first-failure censored order statistics easily. Also, the progressive first-failure censored plan has advantages in terms of reducing the test time, in which more items are used, but only m of $n \times k$ items are failures.

3 Bayes estimation

In this section, we present the posterior densities of the parameters α and β based on progressively first-failure censored data and then obtain the corresponding Bayes estimates of these parameters. To obtain the joint posterior density of α and β, we assume that α and β are independently distributed as gamma(a, b) and gamma(c, d) priors, respectively. Therefore, the prior density functions of α and β becomes

$$\pi_1(\alpha|a, b) = \begin{cases} \dfrac{b^a}{\Gamma(a)} \alpha^{a-1} e^{-b\alpha} & \text{if } \alpha > 0 \\ 0 & \text{if } \alpha \leq 0, \end{cases} \tag{6}$$

$$\pi_2(\beta|c, d) = \begin{cases} \dfrac{d^c}{\Gamma(c)} \beta^{c-1} e^{-d\beta} & \text{if } \beta > 0 \\ 0 & \text{if } \beta \leq 0, \end{cases} \tag{7}$$

The gamma parameters a, b, c and d are all assumed to be positive. When $a = b = 0$ $(c = d = 0)$, we obtain the non-informative priors of α and β.

Let $X^{\mathbf{R}}_{i:m:n:k}$, $i = 1, 2, \ldots, m$, be the progressively first-failure censored order statistics from GP(α, β) the distribution of reparametrized GP, with censoring R. From (3), the likelihood function is given by

$$\ell(data|\alpha, \beta) = Ak^m \alpha^m \prod_{i=1}^{m} \beta^{\alpha k(\mathbf{R}_i+1)} (x_i + \beta)^{-[\alpha k(\mathbf{R}_i+1)+1]}, \tag{8}$$

where A is defined in (5) and X_i is used instead of $X^{\mathbf{R}}_{i:m:n:k}$.

The joint posterior density function of α and β given the data is given by

$$\pi^*(\alpha,\beta|data) = \frac{\ell(data|\alpha,\beta)\pi_1(\alpha|a,b)\pi_2(\beta|c,d)}{\int_0^\infty\int_0^\infty \ell(data|\alpha,\beta)\pi_1(\alpha|a,b)\pi_2(\beta|c,d)d\alpha d\beta}. \tag{9}$$

Therefore, the posterior density function of α and β given the data can be written as

$$\pi^*(\alpha,\beta|data) \propto \alpha^{m+a-1}\beta^{c-1}e^{-d\beta}\left[\prod_{i=1}^{m}\beta^{-\alpha k}(x_i+\beta)^{-1}\right]$$

$$\times \exp\left[-\alpha(b-k\sum_{i=1}^{m}\mathbf{R}_i\log\beta + k\sum_{i=1}^{m}(\mathbf{R}_i+1)\log(x_i+\beta))\right] \tag{10}$$

The posterior density (10) can be rewritten as

$$\pi^*(\alpha,\beta|data) \propto g_1(\alpha|\beta,data)g_2(\beta|data)h(\alpha,\beta|data), \tag{11}$$

here, $g_1(\alpha|\beta,data)$ is a gamma density function with the shape and scale parameters as $(m+a)$ and $b-k\sum_{i=1}^{m}\mathbf{R}_i\log\beta + k\sum_{i=1}^{m}(\mathbf{R}_i+1)\log(x_i+\beta)$, respectively, $g_2(\beta|data)$ is a proper density function given by

$$g_2(\beta|data) \propto \frac{\beta^{c-1}e^{-d\beta}}{\left(b-k\sum_{i=1}^{m}\mathbf{R}_i\log\beta + k\sum_{i=1}^{m}(\mathbf{R}_i+1)\log(x_i+\beta)\right)^{m+a}}. \tag{12}$$

Moreover

$$h(\alpha,\beta|data) = \prod_{i=1}^{m}\beta^{\alpha k}(x_i+\beta)^{-1}. \tag{13}$$

Therefore, the Bayes estimate of any function of α and β, say $g(\alpha,\beta)$ under the squared error loss function is

$$g_B(\alpha,\beta) = \frac{\int_0^\infty\int_0^\infty g(\alpha,\beta)g_1(\alpha|\beta,data)g_2(\beta|data)h(\alpha,\beta|data)d\alpha d\beta}{\int_0^\infty\int_0^\infty g_1(\alpha|\beta,data)g_2(\beta|data)h(\alpha,\beta|data)d\alpha d\beta}. \tag{14}$$

It is not possible to compute (14) analytically. We propose to approximate (14) by using importance sampling technique as suggested by Chen and Shao [27]. The details are explained below.

3.1 Importance sampling

Importance sampling is a useful technique for estimations, now we would like to provide the importance sampling procedure to compute the Bayes estimates for parameters of GP(α,β) the distribution of reparametrized GP, and any function of the parameters say $g(\alpha,\beta)=\theta$.

As mentioned previously that $g_1(\alpha|\beta,data)$ is a gamma density and, therefore, samples of α can be easily generated using any gamma generating routine. However, in our case, the proper density function of β equation (12) cannot be reduced analytically to well known distributions and therefore it is not possible to sample directly by standard methods, but the plot of it (see Figure 1) show that it is similar to normal distribution. So to generate random numbers from this distribution, we use the Metropolis-Hastings method with normal proposal distribution. Using Metropolis-Hastings method, simulation based consistent estimate of $E(g(\alpha,\beta)) = E(\theta)$ can be obtained using Algorithm 1 as follows

Algorithm 1.

Step 1.Start with an $(\alpha^{(0)},\beta^{(0)})$.
Step 2.Set $t=1$.
Step 3.Generate $\beta^{(t)}$ from $g_2(.|data)$ using the method developed by Metropolis et al. [28] with the N$(\beta^{(t-1)},\sigma^2)$ proposal distribution.

Where σ^2 is the variance of β obtained using variance-covariance matrix.

Step 4.Generate $\alpha^{(t)}$ from $g_1(.|\beta^{(t)}, data)$.
Step 5.Compute $\beta^{(t)}$ and $\alpha^{(t)}$.
Step 6.Set $t = t+1$.
Step 7.Repeat Step $3-6$ N times and obtain $(\alpha_1,\beta_1),(\alpha_2,\beta_2),...,(\alpha_N,\beta_N)$.
Step 8.An approximate Bayes estimate of θ under a squared error loss function can be obtained as

$$\hat{g}(\alpha,\beta) = \hat{\theta} = \frac{\frac{1}{N-M}\sum_{i=M+1}^{N} \theta_i h(\alpha_i,\beta_i|data)}{\frac{1}{N-M}\sum_{i=M+1}^{N} h(\alpha_i,\beta_i|data)}.$$

where M is burn-in.
Step 9.Obtain the posterior variance of $\theta = g(\alpha,\beta)$ as

$$\hat{V}(\alpha,\beta|data) = \frac{\frac{1}{N-M}\sum_{i=M+1}^{N} (\theta_i-\hat{\theta})^2 h(\alpha_i,\beta_i|data)}{\frac{1}{N-M}\sum_{i=M+1}^{N} h(\alpha_i,\beta_i|data)}.$$

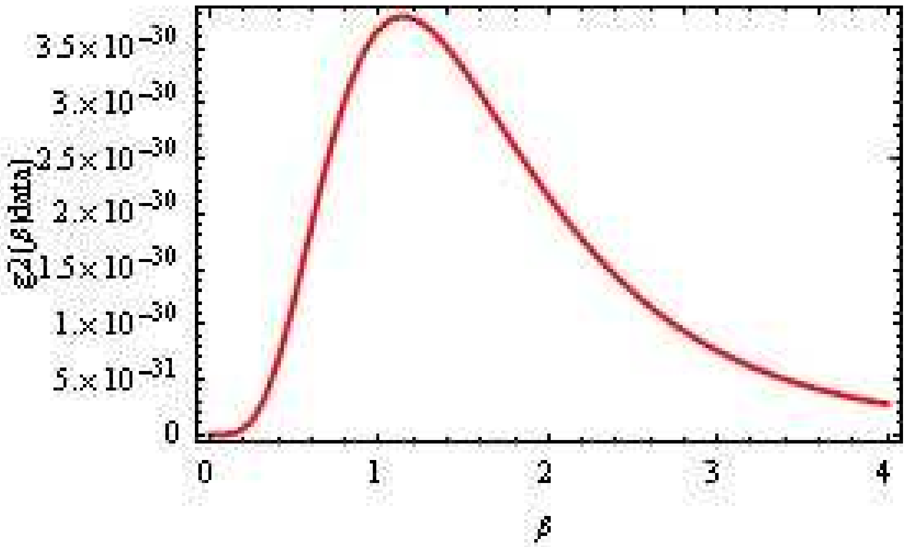

Fig. 1: Posterior density function of β

4 Monte Carlo simulations

In order to compare the proposed Bayes estimators with the MLEs, we simulated 1000 progressively first-failure censored samples from a GP(α,β) the distribution of reparametrized GP. The samples were simulated by using the algorithm described in Balakrishnan and Sandhu (1995). We used a different sample of sizes (n), different effective sample of sizes (m), different k $(k = 1,4)$, different hyperparameters (a,b,c,d), and different of sampling schemes (i.e., different R_i values).We used two sets of parameter values $\alpha = 0.22$, $\beta = 1.5$ and $\alpha = 0.5$, $\beta = 2.2$, mainly to compare the MLEs and different Bayes estimators and also to explore their effects on different parameter values. First, we used the noninformative gamma priors for both the parameters, that is, when the hyperparameters are 0. We call it prior 0: $a = b = c = d = 0$. Note that as the hyperparameters go to 0, the prior density becomes inversely proportional to its argument and also becomes improper. This density is commonly used as an improper prior for parameters in the range of 0 to infinity, and this prior is not specifically related to the gamma density. For computing Bayes estimators, other than prior 0, we also used informative prior, including prior 1, $a = 1$, $b = 1$, $c = 3$ and $d = 2$. In two cases, we used the squared error loss function to compute the Bayes estimates. We also computed the Bayes estimates based on 10,000 MCMC samples and discard the first 1000 values as 'burn-in'

It is clear from Tables 1 and 2 that the proposed Bayes estimators perform very well for different n and m. As expected, the Bayes estimators based on informative prior perform much better than the Bayes estimators based on noninformative prior in terms of biases, MSEs. Also the Bayes estimators based on noninformative prior and informative prior perform much better than the MLEs in terms of biases, MSEs.

Table 1. Average values of the different estimators and the corresponding MSEs it in parentheses when $\alpha = 0.22$ and $\beta = 1.5$.

k	n	m	Scheme	MLE		Bayes (proir 0)		Bayes (prior 1)	
				α	β	α	β	α	β
1	30	20	$(10, 19^0)$	0.2327	1.8364	0.2248	1.7703	0.2349	1.6902
				(0.0043)	(0.4508)	(0.0038)	(0.3423)	(0.0032)	(0.1263)
			$(5^0, 10^1, 5^0)$	0.2366	1.8022	0.2244	1.8064	0.2336	1.6915
				(0.0052)	(0.4641)	(0.0038)	(0.3943)	(0.0033)	(0.1296)
			$(19^0, 10)$	0.2338	1.8074	0.2206	1.8346	0.2289	1.7032
				(0.0057)	(0.4794)	(0.0039)	(0.3999)	(0.0033)	(0.1298)
	40	20	$(20, 19^0)$	0.2373	1.7935	0.2269	1.7971	0.2375	1.7083
				(0.0042)	(0.4408)	(0.0038)	(0.3403)	(0.0031)	(0.1257)
			$(5^0, 10^2, 5^0)$	0.2413	1.7007	0.2343	1.8077	0.2397	1.6879
				(0.0045)	(0.4723)	(0.0040)	(0.3592)	(0.0035)	(0.1262)
			$(19^0, 20)$	0.2526	1.7145	0.2392	1.7093	0.2408	1.7214
				(0.0050)	(0.4837)	(0.0045)	(0.3933)	(0.0036)	(0.1283)
	40	30	$(10, 29^0)$	0.2290	1.8481	0.2214	1.7198	0.2308	1.6594
				(0.0030)	(0.4256)	(0.0029)	(0.3346)	(0.0022)	(0.1208)
			$(10^0, 10^1, 10^0)$	0.2309	1.7627	0.2220	1.7353	0.2288	1.7083
				(0.0031)	(0.4313)	(0.0030)	(0.3549)	(0.0023)	(0.1208)
			$(29^0, 10)$	0.2367	1.7707	0.2291	1.8558	0.2338	1.7060
				(0.0033)	(0.4088)	(0.0030)	(0.3732)	(0.0023)	(0.1401)
4	30	20	$(10, 19^0)$	0.2569	1.7629	0.2106	1.6223	0.1959	1.5548
				(0.0025)	(0.3321)	(0.0023)	(0.3176)	(0.0021)	(0.1114)
			$(5^0, 10^1, 5^0)$	0.2511	1.6563	0.2286	1.4449	0.2082	1.4972
				(0.0026)	(0.3229)	(0.0023)	(0.3223)	(0.0021)	(0.1205)
			$(19^0, 10)$	0.2767	1.6585	0.1974	1.5366	0.2094	1.6022
				(0.0028)	(0.3342)	(0.0024)	(0.3467)	(0.0022)	(0.1228)
	40	20	$(20, 19^0)$	0.2614	1.5897	0.1921	1.5533	0.2034	1.5850
				(0.0022)	(0.3198)	(0.0021)	(0.2963)	(0.0020)	(0.1012)
			$(5^0, 10^2, 5^0)$	0.2545	1.5439	0.1868	1.5385	0.2094	1.6022
				(0.0022)	(0.3226)	(0.0021)	(0.2967)	(0.0021)	(0.1068)
			$(19^0, 20)$	0.2852	1.5232	0.2196	1.4405	0.2099	1.5681
				(0.0023)	(0.3325)	(0.0022)	(0.2975)	(0.0021)	(0.1116)
	40	30	$(10, 29^0)$	0.2567	1.5791	0.2272	1.5669	0.2066	1.5768
				(0.0019)	(0.2664)	(0.0018)	(0.2519)	(0.0017)	(0.0994)
			$(10^0, 10^1, 10^0)$	0.2409	1.6295	0.2117	1.5873	0.1901	1.5849
				(0.0019)	(0.2666)	(0.0018)	(0.2598)	(0.0017)	(0.0996)
			$(29^0, 10)$	0.2587	1.6234	0.2059	1.5227	0.2164	1.5517
				(0.0020)	(0.2668)	(0.0019)	(0.2615)	(0.0018)	(0.0997)

Table 2. Average values of the different estimators and the corresponding MSEs it in parentheses when $\alpha = 0.5$ and $\beta = 2.2$.

k	n	m	Scheme	MLE		Bayes (proir 0)		Bayes (prior 1)	
				α	β	α	β	α	β
1	30	20	$(10, 19^0)$	0.5485 (0.0337)	2.4789 (0.2941)	0.4507 (0.0314)	1.7974 (0.2837)	0.4993 (0.0145)	1.8477 (0.1257)
			$(5^0, 10^1, 5^0)$	0.5271 (0.0344)	2.4346 (0.3087)	0.4508 (0.0331)	2.0163 (0.3051)	0.4626 (0.0155)	1.8399 (0.1381)
			$(19^0, 10)$	0.5076 (0.0353)	2.4846 (0.3109)	0.4210 (0.0332)	1.9355 (0.3064)	0.4296 (0.0159)	1.7781 (0.1395)
	40	20	$(20, 19^0)$	0.5186 (0.0312)	2.4294 (0.2573)	0.4434 (0.0300)	1.9059 (0.2562)	0.4686 (0.0127)	1.8583 (0.1169)
			(20^1)	0.5179 (0.0315)	2.4204 (0.2644)	0.4397 (0.0306)	1.9761 (0.2674)	0.4448 (0.0129)	1.9087 (0.1194)
			$(19^0, 20)$	0.5132 (0.0316)	2.5903 (0.2850)	0.4152 (0.0309)	1.9271 (0.2797)	0.4178 (0.0138)	1.8600 (0.1263)
	40	30	$(10, 29^0)$	0.5194 (0.0305)	2.3885 (0.2167)	0.4110 (0.0293)	1.9243 (0.2097)	0.4693 (0.0084)	1.9751 (0.1099)
			$(10^0, 10^1, 10^0)$	0.5324 (0.0308)	2.5341 (0.2187)	0.4489 (0.0281)	1.9807 (0.2104)	0.4554 (0.0085)	1.9448 (0.1148)
			$(29^0, 10)$	0.5281 (0.0310)	2.4653 (0.2274)	0.4446 (0.0289)	1.9699 (0.2140)	0.4526 (0.0085)	1.8204 (0.1282)
4	30	20	$(10, 19^0)$	0.5656 (0.0277)	2.6291 (0.2001)	0.5966 (0.0262)	1.9932 (0.2023)	0.5487 (0.0081)	1.9629 (0.1072)
			$(5^0, 10^1, 5^0)$	0.5247 (0.0284)	2.5386 (0.2328)	0.5374 (0.0276)	2.1063 (0.2131)	0.5713 (0.0088)	2.1166 (0.1043)
			$(19^0, 10)$	0.6175 (0.0292)	2.3532 (0.2377)	0.5196 (0.0282)	2.3883 (0.2145)	0.5760 (0.0093)	1.9998 (0.1131)
	40	20	$(20, 19^0)$	0.5556 (0.0245)	2.1645 (0.1970)	0.5241 (0.0233)	2.2522 (0.1903)	0.5502 (0.0079)	2.4339 (0.1023)
			(20^1)	0.5692 (0.0256)	2.3340 (0.2006)	0.5270 (0.0236)	2.3109 (0.1992)	0.5883 (0.0083)	2.1262 (0.1033)
			$(19^0, 20)$	0.6120 (0.0265)	2.3269 (0.2083)	0.5334 (0.0243)	1.9992 (0.2054)	0.5914 (0.0089)	2.2128 (0.1126)
	40	30	$(10, 29^0)$	0.4734 (0.0199)	2.1088 (0.1903)	0.5349 (0.0192)	1.9983 (0.1896)	0.5544 (0.0073)	2.2434 (0.0987)
			$(10^0, 10^1, 10^0)$	0.5154 (0.0216)	2.4364 (0.2049)	0.5142 (0.0201)	2.2989 (0.1973)	0.5201 (0.0091)	2.396 (0.0999)
			$(29^0, 10)$	0.5209 (0.0221)	2.2412 (0.2099)	0.5247 (0.0206)	2.3721 (0.1988)	0.5546 (0.0092)	2.5138 (0.1022)

5 Bayesian prediction for future order statistics

Suppose that $X^{\mathbf{R}}_{1:m:n:k}, X^{\mathbf{R}}_{2:m:n:k}, \ldots, X^{\mathbf{R}}_{m:m:n:k}$, is a progressive first-failure censored sample of size m drawn from a population whose pdf is GP(α,β) distribution, defined by (1), and that $Y_1, Y_2, \ldots, Y_{m_1}$ is a second independent random sample (of size m_1) of future observations from the same distribution. Bayesian prediction bounds are obtained for some order statistics of the future observations $Y_1, Y_2, \ldots, Y_{m_1}$. On the other hand, let $X^{\mathbf{R}}_{1:m:n:k}, X^{\mathbf{R}}_{2:m:n:k}, \ldots, X^{\mathbf{R}}_{m:m:n:k}$ and $Y_1, Y_2, \ldots, Y_{m_1}$ represent the informative sample from a random sample of size m, and a future ordered sample of size m_1, respectively. It is further assumed that the two samples are independent and each of their corresponding random samples is obtained from the same distribution function. Our aim is to make Bayesian prediction about the s^{th}, $1 < s < m_1$, ordered lifetime in a future sample of size m_1.

Let Y_s be the s^{th} ordered lifetime in the future sample of size m_1. The density function of Y_s for given α, β is of the form

$$g_{(s)}(y_s|\alpha,\beta) = D(s)\left[1 - F(y_s|\alpha,\beta)\right]^{(m_1-s)}\left[F(y_s|\alpha,\beta)\right]^{s-1} f(y_s|\alpha,\beta), \quad \alpha,\beta > 0, \tag{15}$$

where $D(s) = s\binom{m_1}{s}$.
here $f(.|\alpha,\beta)$ is given in (1) and $F(.|\alpha,\beta)$ denotes the corresponding cumulative distribution function of $f(.|\alpha,\beta)$ as given in (2), substituting (1) and (2) in (15),we obtain

$$g_{(s)}(y_s|\alpha,\beta) = D(s)\alpha\,(y_s+\beta)^{-1}\left[\beta^{\alpha}\,(y_s+\beta)^{-\alpha}\right]^{m_1(s)}\left[1-\beta^{\alpha}\,(y_s+\beta)^{-\alpha}\right]^{s-1}, \tag{16}$$

where $m_1(s) = m_1 - s + 1$.
By using the binomial expansion, the density (16) takes the form

$$g_{(s)}(y_s|\alpha,\beta) = D(s)\alpha\,(y_s+\beta)^{-1}{}_{i=0}^{s-1} a_j(s)\left[\beta^{\alpha}\,(y_s+\beta)^{-\alpha}\right]^{m_{1_j}(s)},\ \ y_s > 0, \tag{17}$$

where

$$a_j(s) = (-1)^j\binom{s-1}{j} \text{ and } m_{1_j}(s) = m_1 - s + j + 1. \tag{18}$$

The Bayes predictive density function of Y_s is given by

$$g^*_{(s)}(y_s|data) = \int_0^{\infty}\int_0^{\infty} g_{(s)}(y_s|\alpha,\beta)\pi^*(\alpha,\beta|data)d\alpha d\beta, \tag{19}$$

where $\pi^*(\alpha,\beta|data)$ is the joint posterior density of α and β as given in (11). It is immediate that $g^*_{(s)}(y_s|data)$ can not be expressed in closed form and hence it can not be evaluated analytically.

A simulation based consistent estimator of $g^*_{(s)}(y_s|data)$, can be obtained by using the Gibbs sampling procedure as described in Section 3. Suppose $\{(\alpha_i,\beta_i), i = 1,2,...,N\}$ are MCMC samples obtained from $\pi^*(\alpha,\beta|data)$, using Gibbs sampling technique, the simulation consistent estimator of $g^*_{(s)}(y_s|data)$, can be obtained as

$$\hat{g}^*_{(s)}(y_s|data) = {}_{i=M+1}^{N}\, g_{(s)}(y_s|\alpha_i,\beta_i)w_i, \tag{20}$$

and a simulation consistent estimator of the predictive distribution of Y_s say $G^*_{(s)}(.|data)$ can be obtained as

$$\hat{G}^*_{(s)}(y_s|data) = {}_{i=M+1}^{N}\, G_{(s)}(y_s|\alpha_i,\beta_i)w_i, \tag{21}$$

where

$$w_i = \frac{h(\alpha_i,\beta_i|data)}{{}_{i=M+1}^{N} h(\alpha_i,\beta_i|data)};\ \ i = M+1,...,N \text{ and } M \text{ is burn-in}, \tag{22}$$

and $G_{(s)}(y_s|\alpha,\beta)$ denotes the distribution function corresponding to the density function $g_{(s)}(y_s|\alpha,\beta)$, here

$$G_{(s)}(y_s|\alpha,\beta) = D(s)\alpha{}_{i=0}^{s-1} a_j(s)\frac{1}{\alpha m_{1_j}(s)}\left[1-\left(\beta^{\alpha}\,(y_s+\beta)^{-\alpha}\right)^{m_{1_j}(s)}\right], \tag{23}$$

where $a_j(s)$ and $m_{1_j}(s)$ are defined in (18). It should be noted that the MCMC samples $\{(\alpha_i,\beta_i), i = 1,2,...,N\}$ can be used to compute $\hat{g}^*_{(s)}(y_s|data)$ or $\hat{G}^*_{(s)}(y_s|data)$ for all Y_s. Moreover, a symmetric $100\gamma\%$ predictive interval for Y_s can be obtained by solving the non-linear equations (24) and (25), for the lower bound, L and upper bound, U

$$\frac{1+\gamma}{2} = P[Y_s > L|data] = 1 - G^*_{(s)}(L|data) \Longrightarrow G^*_{(s)}(L|data) = \frac{1}{2} - \frac{\gamma}{2}, \tag{24}$$

$$\frac{1-\gamma}{2} = P[Y_s > U|data] = 1 - G^*_{(s)}(U|data) \Longrightarrow G^*_{(s)}(U|data) = \frac{1}{2} + \frac{\gamma}{2}. \tag{25}$$

We need to apply a suitable numerical method as they cannot be solved analytically.

6 Bayesian prediction for future record values

Let us consider that $X^{\mathbf{R}}_{1:m:n:k}, X^{\mathbf{R}}_{2:m:n:k}, ..., X^{\mathbf{R}}_{m:m:n:k}$ is a progressive first-failure censored sample of size m with progressive censoring scheme $\mathbf{R} = (R_1, R_2, ..., R_m)$, drawn from a GP$(\alpha, \beta)$ distribution and let $Z_1, Z_2, ..., Z_{m_2}$ is a second independent random sample of size m_2 of future upper record observations drawn from the same population.

The first sample is referred to as the "informative" (past) sample, while the second one is referred to as the (future) sample. Based on an informative progressively first-failure censored sample, our aim is to predict the S^{th} upper record values. The conditional pdf of Z_s for given α, β is given see Ahmadi and MirMostafaee [22], by

$$h_{(s)}\left(z_s|\alpha,\beta\right) = \frac{\left[-\log(1-F\left(z_s|\alpha,\beta\right))\right]^{s-1}}{(s-1)!} f\left(z_s|\alpha,\beta\right), \tag{26}$$

where $F(.|\alpha,\beta)$ and $f(.|\alpha,\beta)$ are given in (2) and (1). Applying (2) and (1) in (26) we obtain

$$h_{(s)}\left(z_s|\alpha,\beta\right) = \frac{1}{(s-1)!}\alpha\beta^{\alpha}(z_s+\beta)^{-(\alpha+1)}[-\log\left(\beta^{\alpha}(z_s+\beta)^{-\alpha}\right)]^{s-1}. \tag{27}$$

The Bayes predictive density function of Y_s is then

$$h^{*}_{(s)}(z_s|data) = \int_0^{\infty}\int_0^{\infty} h_{(s)}(z_s|\alpha,\beta)\pi^{*}(\alpha,\beta|data)d\alpha d\beta, \tag{28}$$

As before, based on MCMC samples $\{(\alpha_i,\beta_i), i = 1,2,...,N\}$, a simulation consistent estimator of $h^{*}_{(s)}(z_s|data)$, can be obtained as

$$\hat{h}^{*}_{(s)}(z_s|data) =_{i=M+1}^{N} h_{(s)}(z_s|\alpha_i,\beta_i)w_i, \tag{29}$$

and a simulation consistent estimator of the predictive distribution of Y_s say $G^{*}_{(s)}(.|data)$ can be obtained as

$$\hat{H}^{*}_{(s)}(z_s|data) =_{i=M+1}^{N} H_{(s)}(z_s|\alpha_i,\beta_i)w_i, \tag{30}$$

where w_i is same as defined in (22) and $H_{(s)}(z_s|\alpha,\beta)$ denotes the distribution function corresponding to the density function $h_{(s)}(z_s|\alpha,\beta)$, we simply obtain

$$\begin{aligned} H_{(s)}(z_s|\alpha,\beta) &= \frac{1}{(s-1)!}\int_0^{z_s}\alpha\beta^{\alpha}(t_s+\beta)^{-(\alpha+1)}[-\log\left(\beta^{\alpha}(t_s+\beta)^{-\alpha}\right)]^{s-1}dt. \\ &= -\frac{1}{(s-1)!}\int_1^{\beta^{\alpha}(z_s+\beta)^{-\alpha}}(-\log(u))^{(s-1)}du. \\ &= \frac{1}{(s-1)!}\left[\Gamma(s)-\Gamma\left(s,-\log\left(\beta^{\alpha}(z_s+\beta)^{-\alpha}\right)\right)\right] tag31 \end{aligned} \tag{15}$$

It should be noted that the MCMC samples $\{(\alpha_i,\beta_i), i = 1,2,...,N\}$ can be used to compute $\hat{h}^{*}_{(s)}(z_s|data)$ or $\hat{H}^{*}_{(s)}(z_s|data)$ for all Z_s. Moreover, a symmetric $100\gamma\%$ predictive interval for Z_s can be obtained by solving the non-linear equations (32) and (33), for the lower bound, L and upper bound, U

$$\frac{1+\gamma}{2} = P[Z_s > L|data] = 1 - H^{*}_{(s)}(L|data) \Longrightarrow H^{*}_{(s)}(L|data) = \frac{1}{2} - \frac{\gamma}{2}, \tag{32}$$

$$\frac{1-\gamma}{2} = P[Z_s > U|data] = 1 - H^{*}_{(s)}(U|data) \Longrightarrow H^{*}_{(s)}(U|data) = \frac{1}{2} + \frac{\gamma}{2}. \tag{33}$$

In this case also it is not possible to obtain the solutions analytically, and one needs a suitable numerical technique for solving these non-linear equations.

7 Illustrative example

To illustrate the methods proposed in the previous sections. A set of data consisting of 60 observations were generated from a GP(α,β) the distribution of reparametrized GP with parameters $(\alpha,\beta) = (0.5, 1.5)$, the generated data are given in Table 3

Table 3. Simulated data from GP$(0.5, 1.5)$.

1.0501	6.7579	1.2581	3.1730	1.3025	0.8422	41.464	6.2518	4.9706	15.896
4.2410	2.7351	70.796	1.9053	3.8841	9.1233	1.3397	215.4	1.1411	42.465
0.1289	7.9931	147.20	6.3203	0.0088	25.356	4.7883	2.7279	7.6829	0.3433
27.402	5.7937	4.1718	1.5419	9.2444	0.1746	5.0694	29.243	1.8672	4.1394
1.5692	4.4158	0.0515	16.594	1.5788	5.6310	0.6328	56.232	1.5865	1.2436
23.018	0.5258	5.8716	9.9957	1.1576	28.908	4.1059	49.807	5.9503	0.3648

To illustrate the use of the estimation methods proposed in this article, we assume that the Simulated data are randomly grouped into 30 groups with ($k = 2$) items within each group. These groups are: {0.1289, 1.5788}, {6.2518, 9.1233}, {0.0515, 1.1576}, {5.6310, 49.807}, {4.7883, 147.20}, {5.8716, 16.594}, {0.8422, 1.3397}, {23.018, 28.908}, {4.1718, 27.402}, {41.464, 215.4}, {0.3648, 4.1059}, {6.3203, 29.243}, {5.7937, 9.9957}, {5.0694, 56.232}, {4.2410, 9.2444}, {7.9931, 25.356}, {3.1730, 5.9503}, {1.1411, 1.5865}, {4.4158, 70.796}, {0.1746, 2.7279}, {1.5692,1.8672}, {4.9706, 42.465}, {2.7351, 3.8841}, {6.7579, 15.896}, {0.3433, 1.2436}, {1.2581, 1.30254}, {4.1394, 7.6829}, {0.0088, 1.0501}, {0.5258, 1.5419}, {0.6328, 1.9053}. Suppose that the pre-determined progressively first-failure censoring plan is applied using progressive censoring scheme $\mathbf{R} = (2,1,1,2,0,0,2,2,0,2,0,2,0,1,0)$. The following progressively first-failure censored data of size ($m = 15$) out of 30 groups were observed: 0.0088, 0.0515, 0.1289, 0.1746, 0.3433, 0.3648, 0.5258, 0.6328, 0.8422, 1.1411, 1.2581, 1.5692, 6.2518, 6.7579, 7.9931.
For this example, 15 groups are censored, and 15 first failure times are observed. Using the progressively first-failure censored sample the MLE's of α and β, are 0.5473 and 1.6811, respectively. we apply the Gibbs and Metropolis samplers with the help of importance sampling technique to determine the Bayesian estimation and prediction intervals, we assumed that both the parameters are unknown. Since we do not have any prior information available, we used noninformative priors $(a = b = c = d = 0)$ on both α and β. The density function of $g_2(\beta|data)$ as given in (12) is plotted Figure (1). It can be approximated by normal distribution function as mentioned in the Subsection 3.1. Now using Algorithm 1, we generate $10,000$ MCMC samples and discard the first 1000 values as 'burn-in', based on them we compute the Bayes estimates of α and β as 0.5186 and 1.51385, respectively. As expected the Bayes estimates under the non-informative prior, and the MLE's are quite close to each other. Moreover, the result of 90% and 95% highest posterior density (HPD) credible intervals of α and β are given in Tables 4 and 5 for the future order statistics and future upper record values, respectively.

Table 4. Two sample prediction for the future order statistics

90% (HPD) credible intervals for Y_S			95% (HPD) credible intervals for Y_S	
Y_S	[Lower,Upper]	Length	[Lower,Upper]	Length
Y_1	[0.0084,0.7250]	0.7166	[0.0042,0.9799]	0.9758
Y_2	[0.0601,1.4470]	1.3869	[0.0402,1.9357]	1.8956
Y_3	[0.1471,2.4408]	2.2938	[0.1075,3.3300]	3.2224
Y_4	[0.2665,3.9269]	3.6603	[0.2032,5.5573]	5.3540
Y_5	[0.4230,6.3025]	5.8795	[0.3303,9.3826]	9.0523

Table 5. Two sample prediction for the future upper record values

90% (HPD) credible intervals for Z_S			95% (HPD) credible intervals for Z_S	
Z_S	[Lower,Upper]	Length	[Lower,Upper]	Length
Z_1	[0.0439,2.6625]	2.6186	[0.0213,2.8424]	2.8212
Z_2	[0.0822,3.2278]	3.1457	[0.0549,5.7681]	5.7132
Z_3	[0.1482,4.9041]	4.7558	[0.1087,6.6587]	6.5500
Z_4	[0.2524,5.3410]	5.0886	[0.2231,7.9281]	7.7050
Z_5	[0.4437,7.5549]	7.1112	[0.3376,10.8731]	10.5355

8 Conclusions

In this paper, Bayesian inference and prediction problems of the generalized Pareto (GP) distribution based on progressive first-failure censored data are obtained for future order statistics and future upper record values. The prior belief of the model is represented by the independent gamma priors on the both scale and shape parameters. The squared error loss function is used. We used Gibbs sampling technique to generate MCMC samples and then using importance sampling methodology we computed the Bayes estimates. The same MCMC samples were used for two sample prediction problems. The details have been explained using a simulated data.

Acknowledgement

The authors would like to express their thanks to the referees for their useful comments and suggestions on the original version of this manuscript.

References

[1] D. Kundu, Technometrics, **50**, 144-154 (2008).
[2] M. Z. Raqab, A. R. Asgharzadeh, R. Valiollahi, Computational Statistics and Data Analysis, **54**, 1732-1743 (2010).
[3] N. Balakrishnan, R. Aggarwala, Birkhauser, Boston, Berlin, (2000).
[4] N. Balakrishnan, Test, **16**, 211–296 (2007).
[5] L. G. Johnson, Elsevier, Amsterdam, (1964).
[6] C. H. Jun, S. Balamurali, S. H. Lee, IEEE Transactions on Reliability, **55**, 53-58 (2006).
[7] J. W. WU, W. L. Hung, C. H. Tsai, Statistics **37**, 517-525 (2003).
[8] J. W. Wu, H. Y. Yu, Applied Mathematics and computation, **163**, 443-482 (2005).
[9] S. J. Wu, C. Kuş, Computational Statistics and Data Analysis, **53**, 3659-3670 (2009).
[10] G. J. Hahn, W. Q. Meeker, John Wiley and Sons, (1991).
[11] A. M. Nigm, Communications in Statistics-Theory and Methods, **17**, 287-297 (1988).
[12] E. K. AL-Huesaini, Z. F. Jaheen, Communications in Statistics-Theory and Methods, **24**, 1829-1842 (1995).
[13] E. K. AL-Huesaini, Z. F. Jaheen, Statistical Planning and Inference, **55**, 23-37 (1996).
[14] M. A. M. Ali Mousa, Z. F. Jaheen, Statistics, **48**, 337-344 (1997).
[15] M. A. M. Ali Mousa, Z. F. Jaheen, Applied Statistical Science, **7**, 103-111 (1998).
[16] Z. F. Jaheen, B.N. AL-Matrafi, Computers and Mathematics with Applications, **44**, 587-594 (2002).
[17] A. H. Abd Ellah, Indian Journal Pure and Applied Mathematics, **34**, 101-109 (2003).
[18] A. H. Abd Ellah, Journal of Statistical Research and Training Center **3**, 139-158 (2006).
[19] A. A. Alamm, M. Z. Raqab, M. T. Madi, JIRSS, **6** , 17-30 (2007).
[20] D. Kundu, H. Howlader, Computational Statistics and Data Analysis, **54**, 1547 1558 (2010).
[21] J. Ahmadi, S. M. T. K. MirMostafaee, N. Balakrishnan, Statistics iFirst, 1–13 (2010).
[22] J. Ahmadi, S. M. T. K. MirMostafaee, Statistics and Probability Letters, **79**, 977-983 (2009).
[23] M. A. M. Ali Mousa, S. A. AL-Sagheer, Communications in Statistics-Theory and Methods, **34**, 2353–2361 (2005).
[24] M. Habibullah, M. Ahsanullah, Communication in Statistics, Theory and Methods, **29**, 1597-1609 (2000).
[25] S. K. Upadhyay, M. Peshwani, Communication in Statistics, Theory and Methods, **32**, 381-405 (2003).
[26] E. K. Al-Hussaini, Statistical Planning and Inference, **79**, 79- 91 (1999).
[27] M-H. Chen, Q-M. Shao, Computational and Graphical Statistics, **8** , 69-92 (1999).
[28] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, Journal Chemical Physics, **21**, 1087–1091 (1953).