J. Statist. Res. Iran **16** (2019): 73–99
DOI: ???

Comparison of Record Ranked Set Sampling and Ordinary Records in Prediction of Future Record Statistics from an Exponential Distribution

Ehsan Golzade Gervi[†], Parviz Nasiri[†,*] and Mahdi Salehi[‡]

[†] Payame Noor University
[‡] University of Neyshabur

Received: 12/13/2019 Approved: 11/8/2020

Abstract. In some situations, considering a suitable sampling scheme, to reduce the cost and increase efficiency is crucial. In this study, based on a record ranked set sampling scheme, the likelihood and Bayesian prediction of upper record values from a future sequence are discussed in the exponential model. To this end, under an upper *record ranked set sample (RRSS)* as an informative sample, the maximum likelihood as well as the Bayes point predictors for future upper record values under squared error (SE) and linear-exponential (LINEX) loss functions are obtained. Furthermore, based on a RRSS scheme, two Bayesian prediction intervals are presented. Prediction intervals are compared in terms of coverage probability and expected length. The results of the RRSS scheme are compared with the one based on ordinary records. Finally, a real data set concerning the daily heat degree is used to evaluate the theoretical results obtained. The results show that in most of the situations, the RRSS scheme performs better.

Keywords. Bayesian prediction; maximum likelihood prediction; record values; record ranked set sampling scheme.
MSC 2010: 62G30; 62G05.

[*] Corresponding author

1 Introduction

There are experiments where have been done sequentially, and only record data are observed. These types of data arise in a wide variety of practical situations such as hydrology, sports achievements, quality control, the strength of materials, and climatology. To this end, we leave a brief description of record values here.

Let $\{X_i, i \geq 1\}$ be an infinite sequence of independent and identically distributed (i.i.d) random variables with absolutely continuous cumulative distribution function (cdf) $F(x;\theta)$ and corresponding probability density function (pdf) $f(x;\theta)$ where θ is possibly a vector real-valued parameter. Define $Z_n = \max\{X_1, \cdots, X_n\}$ for $n \geq 1$; then an observation X_j is called an *upper record value* if $X_j > Z_{j-1}, \quad j > 1$. An analogous definition can be given for *lower record value.* Some key references are Ahsanullah (1995), Arnold et al. (1998), Nevzorov (2001) and Gulati and Padgett (2003). Predicting future record statistics is one of the concepts in real-life situations. The purpose of statistical prediction is to infer the values of future statistics based on available observations. For more details see, Aitchison and Dunsmore (1975) and Geisser (1993). In many experiments, using sampling methods based on different suitable schemes to reduce the cost and increase efficiency is crucial. The record ranked set sample (RRSS) scheme, as an alternative method for generating record-breaking data, has been formally proposed by Salehi and Ahmadi (2014). Among the authors who worked on this scheme, Salehi and Ahmadi (2015) considered the estimation of stress and strength using upper RRSS from the exponential distribution. They also, with the collaboration of Dey (2016), made a comparison between RRSS scheme and the ordinary record statistics in estimating the unknown parameter of the proportional hazard rate model. They showed that the RRSS scheme outperforms the ordinary record statistics in the frequentist/Bayesian point and interval estimation under that family of distributions. Eskandarzadeh et al. (2016) obtained information measures for the RRSS scheme. Paul and Thomas (2017) proposed concomitant RRSS for situations that measuring of the variable of interest is costly or even impossible. Safaryian et al. (2019) proposed some improved estimators, including the preliminary test estimator, as well as a stein-type shrinkage estimator for stress-strength reliability using record ranked set sampling scheme. Recently, Sadeghpour et al. (2020) considered the estimation of stress and strength reliability using a lower record ranked set sampling scheme under the generalized exponential distribution. Now, to introduce upper RRSS, suppose we have n independent random sequences where the ith sequence sampling is stopped whenever the ith upper record is observed. The only observations available for analysis are the last upper record value in each sequence. This process is called, record ranked set sampling scheme or **Plan B** (in this paper) because it is designed based on the plan of RSS defined by McIntyer (1952). Let us denote the last upper record for the ith sequence in **Plan B** by $R_{i,i}$, then $\mathbf{R_B} = (R_{1,1}, R_{2,2}, \cdots, R_{n,n})'$ will be an upper RRSS of size n. In this study, " $'$ " represents the transpose of a vector. The following observational procedure illustrates this plan

$$\begin{array}{lllll} 1: \ \underline{R_{(1)1}} & & & & \to R_{1,1} = R_{(1)1} \\ 2: \ R_{(1)2} & \underline{R_{(2)2}} & & & \to R_{2,2} = R_{(2)2} \\ \quad \vdots & \quad \vdots & \quad \vdots & & \quad \vdots \quad \vdots \\ n: \ R_{(1)n} & R_{(2)n} & \cdots & \underline{R_{(n)n}} & \to R_{n,n} = R_{(n)n} \end{array} \tag{1}$$

where $R_{(i)j}$ is the ith ordinary (upper) record in the jth sequence. Notice that unlike the ordinary records, here $R_{i,i}$'s are independent random variables and not necessarily ordered with probability one (however, Salehi and Ahmadi (2014) showed that they are ordered with a probability greater than 0.5). Thus, if $R_{i,i}$'s are upper records, then by using the marginal density of ordinary records, the joint pdf of them readily follows as (see, Arnold et al. (1998))

$$f_{\mathbf{R_B}}(\mathbf{r_B}; \theta) = \prod_{i=1}^{n} \frac{\{-\log \bar{F}(r_{i,i}; \theta)\}^{i-1}}{(i-1)!} f(r_{i,i}; \theta), \quad \theta \in \Theta, \tag{2}$$

where $\bar{F}(.) = 1 - F(.)$, $\mathbf{r_B} = (r_{1,1}, r_{2,2}, \cdots, r_{n,n})'$ is the observed vector of $\mathbf{R_B}$ and Θ is the parameter space. In this study, we consider two-sample prediction problems. In **Plan B**, a specific statistic of a future sequence is predicted based on an observed sample. These two samples are usually supposed to be independent and called *future samples* and *informative samples*, respectively. More precisely, we intend to predict ordinary upper record statistics arising from a future sample based on an observed upper RRSS, which is independent of the future sample. Now let us denote the well known inverse sampling scheme in ordinary record statistics by **Plan A** (in this paper). In **Plan A**, there is only one sample in which it is first part is observed and plays the role of informative sample and its second part which is not yet observed plays the role of future a sample. As mentioned earlier, Salehi et al. (2016) compared the Plans **A** and **B** in some issues of inferential statistics including Fisher information and point/interval estimation under both frequentist and Bayesian approaches when the population follows the proportional hazard rate model (including the exponential one). They showed that in most of the situations **Plan B** provides more efficient estimators than their counterparts obtained based on **Plan A**. Many authors have studied on the two-sample prediction. Raqab and Balakrishnan (2008) studied prediction intervals for future records, MirMostafaee and Ahmadi (2011) considered point prediction of future order statistics from an exponential distribution, Ahmadi et al. (2012) obtained Bayesian predictors of k-record values based on progressively censored data from an exponential distribution, Salehi et al. (2015) considered the prediction of order statistics and record values based on an ordered ranked set sample. In this study, after predicting the future record values based on **Plan B** through the Bayesian procedure as well as the maximum likelihood one, we are interested in comparing the performance of these plans on the mentioned issue. Throughout this study, we suppose that the independent random sequence $\{X_i, i = 1, 2, \cdots\}$ comes from the exponential distribution with the pdf

$$f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}} \ , \quad x > 0,\ \theta > 0, \tag{3}$$

and the cdf

$$F(x;\theta) = 1 - e^{-\frac{x}{\theta}}. \tag{4}$$

Thus, the rest of this study is arranged as follows. In Section 2, we derive Bayes point predictors under SE and LINEX loss functions as well as two Bayesian prediction intervals for upper future record values based on RRSS. The maximum likelihood predictor based on RRSS is studied in Section 3. Also, Section 4 contains a brief introduction to inverse sampling scheme and some corresponding results. The comparative study is presented in Section 5 to compare the performance of the proposed predictors. In Section 6, a real data set concerning the daily heat degree, is analyzed for comparing our results based on RRSS with the corresponding one of the inverse sampling scheme. Section 7 concludes.

2 Bayesian Prediction based on Plan B

Suppose $\mathbf{r_B} = (r_{1,1}, \cdots, r_{n,n})'$ be the observation of random vector $\mathbf{R_B} = (R_{1,1}, \cdots, R_{n,n})'$, an upper RRSS of size n from (3). By using (2), the likelihood function of θ given $\mathbf{r_B}$ can be obtained as

$$\begin{aligned} L(\theta|\mathbf{r_B}) &= \prod_{i=1}^{n} \frac{\big(-\log e^{-\frac{r_{i,i}}{\theta}}\big)^{i-1}}{(i-1)!} \frac{1}{\theta} e^{-\frac{r_{i,i}}{\theta}}, \\ &\propto \theta^{-N} e^{-\frac{t}{\theta}}, \end{aligned} \tag{5}$$

where $N = \frac{n(n+1)}{2}$ and $t = \sum_{i=1}^{n} r_{i,i}$ is the observed value of $T = \sum_{i=1}^{n} R_{i,i}$. The random variable $\sum_{i=1}^{n} R_{i,i}$ is a complete sufficient statistic for θ. In the Bayesian approach, we need to specify a prior distribution for the parameter θ. The Inverse-Gamma distribution with pdf

$$\pi(\theta) = \frac{1}{\Gamma(\alpha)} \beta^{-\alpha} \theta^{-\alpha-1} e^{-\frac{1}{\beta\theta}}, \qquad \theta > 0, \tag{6}$$

where $\alpha(>0)$ and $\beta(>0)$ are hyperparameters and $\Gamma(.)$ stands for the complete gamma function makes a conjugate prior for θ. If we denote this distribution by $\theta \sim Inv - Gamma(\alpha, \beta)$ then by using (5) and (6) we get the posterior distribution of θ given $\mathbf{r_B}$ as follows

$$\Pi(\theta|\mathbf{r_B}) = \frac{1}{\Gamma(N+\alpha)(\frac{\beta}{1+\beta t})^{N+\alpha}} \theta^{-N-\alpha-1} e^{-\frac{1}{\theta}(\frac{1}{\beta}+t)}. \tag{7}$$

This implies that, $\theta|\mathbf{R_B} \sim Inv - Gamma(N+\alpha, \frac{\beta}{1+\beta T})$. In the following, we present some point predictors and Bayesian prediction intervals for upper record values in a future sequence based on **Plan B**. First, we present Bayes point predictors.

2.1 Bayes Point Predictor based on Plan B

Suppose that Y_j is jth upper record value from a future sequence. For Bayesian point predicting, we need to obtain the Bayesian predictive density function of Y_j given $\mathbf{r_B}$. The pdf of Y_j can be obtained as follows (see, Arnold et al. (1998))

$$h_{Y_j}(y|\theta) = \frac{1}{\Gamma(j)}\theta^{-j}y^{j-1}e^{-\frac{y}{\theta}}, \qquad y > 0. \tag{8}$$

Thus the predictive density function of Y_j given $\mathbf{r_B}$ is given by (see, Arnold et al.(1998))

$$h^*_{Y_j}(y|\mathbf{r_B}) = \int_\theta h_{Y_j}(y|\theta)\Pi(\theta|\mathbf{r_B})d\theta. \tag{9}$$

By substituting (7) and (8) in (9) and then integrating out the parameter θ, the Bayesian predictive density function of Y_j given $\mathbf{r_B}$, will be obtained as

$$h^*_{Y_j}(y|\mathbf{r_B}) = \frac{1}{B(j, N+\alpha)y}p(y)^j(1-p(y))^{N+\alpha}, \quad y > 0, \tag{10}$$

where

$$p(y) = \frac{\beta y}{1+\beta y+\beta t}.$$

Moreover, $B(.,.)$ is the complete beta function. Now, using (10), we compute Bayes point predictors for Y_j based on an upper record value under the SE and LINEX loss functions. These loss functions are defined respectively as follows

$$L_1(Y_j, \hat{Y}_j) = (\hat{Y}_j - Y_j)^2,$$

$$L_2(Y_j, \hat{Y}_j) = b[e^{a(\hat{Y}_j - Y_j)} - a(\hat{Y}_j - Y_j) - 1], \quad a \neq 0,$$

where $b > 0$ is the scale parameter of the LINEX loss function and without loss of generality, it can be assumed that $b = 1$. The SE loss function is a symmetric loss function. The symmetric nature of this function gives equal weight to overestimation as well as underestimation while in the LINEX loss function, the sign and magnitude of the shape parameter $a \neq 0$ represents the direction and degree of symmetry, respectively. A positive value of a is used when the overestimation is more serious than an underestimation while a negative value of a is vice-versa. For a close to zero, this loss function is approximately SE and therefore almost symmetric (see, Varian (1975) and Zellner (1986)). The Bayes point predictor for Y_j under a symmetric loss as the SE loss function is of the form

$$\hat{Y}_j^{(BS)} = E(Y_j|\mathbf{R_B}) = \frac{j}{N+\alpha-1}(\frac{1}{\beta}+T). \tag{11}$$

As mentioned earlier in the introduction section, T and Y_j are independent random variables. By applying some computations, we get the mean squared prediction error of $\hat{Y}_j^{(BS)}$ under SE loss function (let us denote here by $MSPE(\hat{Y}_j^{(BS)}, Y_j)$ which is the mean of $(\hat{Y}_j^{(BS)} - Y_j)^2$'s) as follows

$$\begin{aligned} MSPE(\hat{Y}_j^{(BS)}, Y_j) = E\big(\hat{Y}_j^{(BS)} - Y_j\big)^2 &= \frac{j^2}{(N+\alpha-1)^2}\big(N\theta^2 + (\frac{1}{\beta}+N\theta)^2\big) \\ &\quad - \frac{2j^2\theta}{N+\alpha-1}(\frac{1}{\beta}+N\theta) + j\theta^2 + j^2\theta^2. \end{aligned} \tag{12}$$

A useful alternative to the SE loss function is the LINEX loss function. The LINEX loss function is a convex but asymmetric loss function. The Bayes point predictor for Y_j under the LINEX loss function is $-\frac{1}{a}\ln[E(e^{-aY_j} \mid \mathbf{r_B})]$. So, from (10), we have

$$\begin{aligned} \hat{Y}_j^{(BL)} &= -\frac{1}{a}\ln\int_0^\infty e^{-ay} h^*_{Y_j}(y|\mathbf{r_B})dy \qquad (13) \\ &= -\frac{1}{a}\ln\int_0^\infty \sum_{k=0}^\infty \frac{(-ay)^k}{k!} h^*_{Y_j}(y|\mathbf{r_B})dy \\ &= -\frac{1}{a}\ln\sum_{k=0}^\infty \frac{\Gamma(k+j)}{\Gamma(j)}\frac{\Gamma(N+\alpha-k)}{\Gamma(N+\alpha)}\frac{(-a(\frac{1}{\beta}+t))^k}{k!}. \end{aligned}$$

There is no closed-form for the predicted risk of $\hat{Y}_j^{(BL)}$, (let us denote here by $PR(\hat{Y}_j^{(BL)}, Y_j)$ which is the mean of $b[e^{a(\hat{Y}_j^{(BL)}-Y_j)} - a(\hat{Y}_j^{(BL)} - Y_j) - 1]$'s). Monte Carlo simulation has to be used here. Moreover, we can see:
(i). Under the modified SE loss function of the form $L_3(Y_j, \hat{Y}_j) = (\frac{\hat{Y}_j}{Y_j} - 1)^2$, the Bayes point predictor for Y_j, is $\frac{E(\frac{1}{Y_j}|\mathbf{r_B})}{E(\frac{1}{Y_j^2}|\mathbf{r_B})}$. So

$$\hat{Y}_j^{(BM)} = \frac{j-2}{N+\alpha+1}(\frac{1}{\beta}+T), \quad j>2.$$

(ii). Under the weighted SE loss function of the form $L_4(Y_j, \hat{Y}_j) = \frac{(\hat{Y}_j - Y_j)^2}{Y_j}$, the Bayes point predictor for Y_j, is $\frac{1}{E(\frac{1}{Y_j}|\mathbf{r_B})}$. Thus

$$\hat{Y}_j^{(BW)} = \frac{j-1}{N+\alpha}(\frac{1}{\beta}+T), \quad j>1.$$

(iii). Under the zero-one loss function which is mostly used in the hypothesis test, the Bayes point predictor for Y_j is the predictive mode which is simply given by

$$\hat{Y}_j^{(BZ)} = \frac{j-1}{N+\alpha+1}(\frac{1}{\beta}+T), \quad j>1.$$

2.2 Bayesian Prediction Interval based on Plan B

Since the point predictors do not tell us much about the precision of the point prediction, two Bayesian prediction intervals for Y_j are constructed. We start with the survival method.

• **Survival method**
To construct a Bayesian prediction interval under survival method, from (10) we have

$$\bar{H}^*_{Y_j}(y|\mathbf{r_B}) = \int_y^\infty \frac{1}{B(j, N+\alpha)z} p(z)^j (1-p(z))^{N+\alpha} dz, \tag{14}$$

where $\bar{H}^*_{Y_j}(y|\mathbf{r_B})$ indicates the Bayesian predictive survival function for Y_j. The Bayesian predictive intervals of a two-sided equi-tailed $100(1-\gamma)\%$ interval for Y_j, can be obtained by solving the following equation

$$P(L(\mathbf{r_B}) < Y_j < U(\mathbf{r_B})) = 1-\gamma. \tag{15}$$

This last statement is equivalent to

$$\bar{H}^*_{Y_j}(L(\mathbf{r_B})) = 1-\frac{\gamma}{2}, \quad \bar{H}^*_{Y_j}(U(\mathbf{r_B})) = \frac{\gamma}{2}, \tag{16}$$

where $L(\mathbf{r_B})$ and $U(\mathbf{r_B})$ denote the lower and upper bounds, respectively. From (14) and (16), we have

$$\int_0^{L(\mathbf{r_B})} \frac{1}{B(j,N+\alpha)y} p(y)^j (1-p(y))^{N+\alpha} dy = \frac{\gamma}{2}.$$

By taking $v = p(y)$, we conclude

$$\int_0^{(1+\frac{\frac{1}{\beta}+t}{L(\mathbf{r_B})})^{-1}} \frac{1}{B(j,N+\alpha)} v^{j-1}(1-v)^{N+\alpha-1} dv = \frac{\gamma}{2}. \qquad (17)$$

Now, let $B_\gamma(c_1,c_2)$ denote the γth quantile of beta distribution with c_1 and c_2 parameters, i.e. $P(\omega \geq B_\gamma(c_1,c_2)) = \gamma$, where $\omega \sim Beta(c_1,c_2)$. Then

$$L(\mathbf{r_B}) = \frac{B_{1-\frac{\gamma}{2}}(j,N+\alpha)}{1-B_{1-\frac{\gamma}{2}}(j,N+\alpha)}(\frac{1}{\beta}+t).$$

Similarly, $U(\mathbf{r_B}) = \frac{B_{\frac{\gamma}{2}}(j,N+\alpha)}{1-B_{\frac{\gamma}{2}}(j,N+\alpha)}(\frac{1}{\beta}+t)$. Therefore, we construct the following $100(1-\gamma)\%$ Bayesian prediction interval for Y_j

$$\Bigg(\frac{B_{1-\frac{\gamma}{2}}(j,N+\alpha)}{1-B_{1-\frac{\gamma}{2}}(j,N+\alpha)}\ (\frac{1}{\beta}+T), \frac{B_{\frac{\gamma}{2}}(j,N+\alpha)}{1-B_{\frac{\gamma}{2}}(j,N+\alpha)}\ (\frac{1}{\beta}+T)\Bigg). \qquad (18)$$

The length of this interval is a random variable. The expected length (EL) of the Bayesian prediction interval in (18) equals to

$$EL_j(B) = (\frac{1}{\beta}+N\theta)\left\{\frac{B_{\frac{\gamma}{2}}(j,N+\alpha) - B_{1-\frac{\gamma}{2}}(j,N+\alpha)}{(1-B_{1-\frac{\gamma}{2}}(j,N+\alpha))(1-B_{\frac{\gamma}{2}}(j,N+\alpha))}\right\}. \qquad (19)$$

Remark 1. By putting $j=1$, from (14) it is clearly concluded that

$$\bar{H}^*_{Y_1}\big(y|\mathbf{r_B}\big) = \left(\frac{\frac{1}{\beta}+t}{y+\frac{1}{\beta}+t}\right)^{N+\alpha}, \quad y>0. \qquad (20)$$

Hence, by substituting (20) into (16), a $100(1-\gamma)\%$ Bayesian prediction interval for Y_1 is given by

$$\left((\frac{1}{\beta}+T)\left\{(1-\frac{\gamma}{2})^{-\frac{1}{N+\alpha}}-1\right\},(\frac{1}{\beta}+T)\left\{(\frac{\gamma}{2})^{-\frac{1}{N+\alpha}}-1\right\}\right). \qquad (21)$$

The EL of the Bayesian prediction interval for Y_1 equals

$$EL_1(B)=(\frac{1}{\beta}+N\theta)\left\{(\frac{\gamma}{2})^{-\frac{1}{N+\alpha}}-(1-\frac{\gamma}{2})^{-\frac{1}{N+\alpha}}\right\}. \qquad (22)$$

• **HPD method**

Another method of constructing Bayesian prediction interval is called the highest posterior density (HPD) method. Now, we construct a HPD interval for Y_j by using (10). Note that, $h^*_{Y_j}(y|\mathbf{r_B})$ is continuous and uni-modal pdf. Thus, we need to solve the following equations:

$$\int_{\zeta_1}^{\zeta_2}\frac{1}{B(j,N+\alpha)y}p(y)^j(1-p(y))^{N+\alpha}dy=1-\gamma\ ,\quad 0<\gamma<1, \qquad (23)$$

and

$$(\frac{\zeta_1}{\zeta_2})^{j-1}=(\frac{\frac{1}{\beta}+t+\zeta_1}{\frac{1}{\beta}+t+\zeta_2})^{N+\alpha+j}. \qquad (24)$$

The interval (ζ_1,ζ_2) is called a Bayesian prediction interval under HPD method for Y_j. It seems that (23) and (24) does not have a closed-form. Monte Carlo simulation have to be utilized here. The corresponding simulation results presented in Table 2. (For more information see, Casella and Berger (2002), Chen and Shao (1999)).

Remark 2. Now, we consider the case where $j = 1$. In this case, the Bayesian predictive density function in (10) for y is strictly decreasing function. It follows that $100(1-\gamma)\%$ Bayesian prediction interval for Y_1 based on the HPD method, is given by

$$\left(0,(\frac{1}{\beta}+T)\{\gamma^{-\frac{1}{N+\alpha}}-1\}\right). \qquad (25)$$

So a two-sided HPD prediction interval cannot be constructed.

3 Maximum Likelihood Prediction based on Plan B

We study maximum likelihood predictor (MLP) of upper record values from a future sequence based on **Plan B**. The principle of maximum likelihood applied to the joint prediction and estimation of a record statistics and an unknown parameter. (see, Kaminsky and Rhodin, (1985)). Here, we assume that the $\mathbf{r_B} = (r_{1,1}, \cdots, r_{n,n})'$ is observed from a population with unknown parameter θ. The predictor of Y_j in a future sequence based on **Plan B** will be obtained through a maximum likelihood approach. The joint pdf of $\mathbf{R_B}$ and Y_j is the product of their marginal pdfs as they are independent, and is

$$f_{R_{1,1},\cdots,R_{n,n},Y_j}(r_{1,1},\cdots,r_{n,n},y;\theta) = \prod_{i=1}^{n} \frac{(-\log e^{-\frac{r_{i,i}}{\theta}})^{i-1}}{(i-1)!}\theta^{-1} \times e^{-\frac{r_{i,i}}{\theta}} \frac{1}{\Gamma(j)}\theta^{-j} y^{j-1} e^{-\frac{y}{\theta}}. \tag{26}$$

Thus

$$\ell \propto -N\log\theta - \frac{t}{\theta} - j\log\theta + (j-1)\log y - \frac{y}{\theta}, \tag{27}$$

where ℓ stands for the log-likelihood function. By differentiating (27) with respect to θ and then y, and equating them to zero, we obtain

$$\hat{Y}_j^P(B) = (j-1)\hat{\theta}_P(B), \quad j>1, \tag{28}$$

where

$$\hat{\theta}_P(B) = \frac{T}{N+1}. \tag{29}$$

The predictor $\hat{Y}_j^P(B)$ is said to be MLP of Y_j, and the estimator $\hat{\theta}_P(B)$ is said to be the predictive maximum likelihood estimator (PMLE) of θ. Since $Y_j \sim Gamma(j,\theta)$ and $T \sim Gamma(N,\theta)$, and moreover T is independent of Y_j, we can compute the MSPE of $\hat{Y}_j^P(B)$ as below

$$MSPE(\hat{Y}_j^P(B), Y_j) = \frac{N\theta^2}{N+1}(1-j^2) + j\theta^2 + j^2\theta^2. \tag{30}$$

Also, we see that, the MSPE of $\hat{Y}_j^P(B)$ is increasing in j when others are kept fixed.

- By putting $j = 2$, $MSPE(\hat{Y}_2^P(B), Y_2) = \dfrac{3\theta^2}{N+1}(N+2)$.

4 Bayesian Prediction based on Plan A

This subsection provides some corresponding results for the inverse sampling scheme. The experiments are done sequentially, the only observations available for analysis are the upper record values, and the sampling is finished when the mth record is observed. This process, is called inverse sampling scheme or **Plan A** (in this paper). Let us denote the ith upper record value by R_i. Contrary to **Plan B**, here R_i's in **Plan A** are not independent random variables. Then, their joint density is given by (see, Arnold et al. (1998))

$$f_{\mathbf{R_A}}(\mathbf{r_A};\theta) = f(r_m;\theta)\prod_{i=1}^{m-1}\frac{f(r_i;\theta)}{\bar{F}(r_i;\theta)}, \quad r_1 < r_2 < \cdots < r_m, \tag{31}$$

where $\mathbf{r_A} = (r_1, \cdots, r_m)'$ is the observed value of $\mathbf{R_A} = (R_1, \cdots, R_m)'$. Hence from (6) and (31) the likelihood function and posterior distribution are, respectively, derived as

$$L(\theta|\mathbf{r_A}) = \theta^{-m}e^{-\frac{r_m}{\theta}},$$

and

$$\Pi(\theta|\mathbf{r_A}) = \frac{1}{\Gamma(m+\alpha)(\frac{\beta}{1+\beta r_m})^{m+\alpha}}\theta^{-m-\alpha-1}e^{-\frac{1}{\theta}(\frac{1}{\beta}+r_m)}. \tag{32}$$

In the following, we discuss the Bayesian prediction problem based on the observed record value of **Plan A**. We first focus on the Bayes point predictor. For predicting of the jth future upper record value, R_j, $j \geq 1$, we need to know the Bayesian predictive density function of R_j given $\mathbf{r_A}$.

4.1 Bayesian Point Prediction based on Plan A

Suppose that, $Y = R_j$, be the jth future upper record value. The pdf of Y given θ can be defined as (see, Arnold et al. (1998))

$$g(y|\theta) = \frac{1}{(j-1)!}[-\log(1-F(y))]^{j-1} f(y;\theta), \quad 0 < y < \infty. \tag{33}$$

From (3) and (4), the pdf of Y given θ is provided by

$$g(y|\theta) = \frac{1}{(j-1)!}\theta^{-j} y^{j-1} e^{-\frac{y}{\theta}}, \qquad 0 < y < \infty. \tag{34}$$

The Bayesian predictive density function of Y given $\mathbf{r_A}$ can be defined as (see, Arnold et al. (1998), p.162)

$$g^*(y|\mathbf{r_A}) = \int_\Theta g(y|\theta)\Pi(\theta|\mathbf{r_A})d\theta \tag{35}$$

Using (32) and (34) relations and integrating out the parameter θ, we have

$$g^*(y|\mathbf{r_A}) = \frac{(\frac{1}{\beta} + y + r_m)^{-(j+\alpha+m)}}{B(j, m+\alpha)}(\frac{1}{\beta} + r_m)^{m+\alpha} y^{j-1}, \tag{36}$$

It is notable that, using (36), we can now obtain the Bayes point predictor of Y under the loss functions mentioned earlier. The Bayes point predictor of R_j under SE loss function is

$$\begin{aligned} \hat{R}_j^{(AS)} &= \int_0^\infty y g^*(y|\mathbf{r_A})dy, \\ &= \frac{j}{m+\alpha-1}(\frac{1}{\beta} + r_m), \end{aligned} \tag{37}$$

and, the MSPE of $\hat{R}_j^{(AS)}$ will be obtained as

$$\begin{aligned} MSPE\big(\hat{R}_j^{(AS)}, R_j\big) = &\big(\frac{j}{m+\alpha-1}\big)^2\big(m\theta^2 + (\frac{1}{\beta} + m\theta)^2\big) \\ &- \frac{2j^2\theta}{m+\alpha-1}\big(\frac{1}{\beta} + m\theta\big) + \theta^2 j(j+1), \end{aligned} \tag{38}$$

where '$\hat{R}_j^{(AS)}$'denotes the Bayes point predictor of the jth future upper record, R_j, under SE loss function based on **Plan A**. The Bayes point predictor of R_j under LINEX loss function, is

$$\hat{R}_j^{(AL)} = -\frac{1}{a} \ln \int_0^{\infty} e^{-ay} g^*(y|\mathbf{r_A}) dy, \tag{39}$$

It can be solved by a simulation. The simulation results presented in Table1.

4.2 Bayesian Prediction Interval based on Plan A

As in estimation, a prediction can be either a point or an interval prediction. We will present two Bayesian prediction intervals based on **Plan A** for R_j.

• **Survival method**

Similar to the calculation of (18) based on **Plan B**, it can be shown that

$$\left(\frac{B_{1-\frac{\gamma}{2}}(j, m+\alpha)}{1 - B_{1-\frac{\gamma}{2}}(j, m+\alpha)} \left(\frac{1}{\beta} + r_m\right), \frac{B_{\frac{\gamma}{2}}(j, m+\alpha)}{1 - B_{\frac{\gamma}{2}}(j, m+\alpha)} \left(\frac{1}{\beta} + r_m\right) \right), \tag{40}$$

is a $100(1-\gamma)\%$ Bayesian prediction interval for R_j based on **Plan A**. The interval length is a function of r_m. Let $EL_j(A)$ denote the corresponding expected length of the Bayesian prediction interval given by (40), then

$$EL_j(A) = \left(\frac{1}{\beta} + m\theta\right) \left\{ \frac{B_{\frac{\gamma}{2}}(j, m+\alpha) - B_{1-\frac{\gamma}{2}}(j, m+\alpha)}{(1 - B_{1-\frac{\gamma}{2}}(j, m+\alpha))(1 - B_{\frac{\gamma}{2}}(j, m+\alpha))} \right\}. \tag{41}$$

• **HPD method**

Using the fact that $g^*_{R_j}(y|\mathbf{r_A})$ is a continuous and uni-modal pdf. In order to find the Bayesian prediction interval under HPD method by (36), the constants ξ_1 and ξ_2 must be found such that they satisfy (42) and (43) simultaneously:

$$\int_{\xi_1}^{\xi_2} g^*(y|\mathbf{r_A}) dy = 1 - \gamma \ , \ 0 < \gamma < 1, \tag{42}$$

$$g^*(\xi_1|\mathbf{r_A}) = g^*(\xi_2|\mathbf{r_A}). \tag{43}$$

The interval (ξ_1, ξ_2) is called a Bayesian prediction interval under HPD method for R_j based on **Plan A**. The equations (42) and (43) does not have a closed-form, thus must be solved by suitable numerical method or Monte Carlo simulation. The corresponding simulation results presented in Table 2. With a little care, we can see that, substituting (m, r_m) instead of (N, T), respectively, yields the predictors based on **Plan A**.

5 Comparison

We compare the predictors obtained under the RE and the REL criteria, which are defined as follows

$$RE = \frac{\text{The MSPE of point predictor based on } \textbf{Plan B}}{\text{The MSPE of point predictor based on } \textbf{Plan A}}, \tag{44}$$

and

$$\begin{cases} REL = \frac{\text{The EL of prediction interval based on } \textbf{Plan B}}{\text{The EL of prediction interval based on } \textbf{Plan A}}, \\ with \\ CP = \frac{1}{K}\sum_{i=1}^{K} I(P_i \leq Y_j \leq Q_i), \end{cases} \tag{45}$$

where $I(.)$ is the indicator function. As mentioned earlier in the subsection 2.1, we use PR instead of MSPE, where PR indicates the predicted risk of point predictor under LINEX loss function. The acronyms EL and CP stand for the average of the expected length and coverage probability for K iterations, respectively. Let (P, Q) be a Bayesian prediction interval for Y_j. Moreover, (P_i, Q_i), $i = 1, ..., K$, observed values of lower and upper bound of the proposed prediction interval. Thus

$$EL = \frac{1}{K}\sum_{i=1}^{K}(Q_i - P_i) \qquad \& \qquad CP = \frac{1}{K}\sum_{i=1}^{K} I(P_i \leq Y_j \leq Q_i).$$

In the following for instance, with a similar manner in (28), it can be shown that

$$\hat{R}_j^P(A) = \frac{j-1}{m+1} r_m. \tag{46}$$

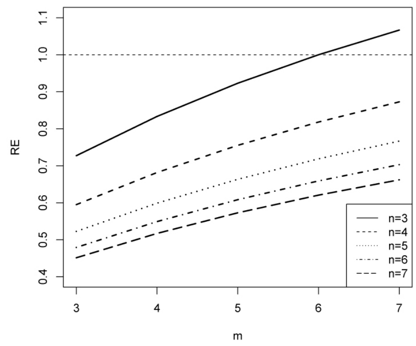

Figure 1. The plot of $RE = \frac{MSPE(\hat{Y}_j^P(B),Y_j)}{MSPE(\hat{R}_j^P(A),R_j)}$ versus m for $j = 8$, $m = 3(1)7$ and $N = 6, 10, 15, 21, 28$.

and

$$MSPE(\hat{R}_j^P(A), R_j) = \frac{m\theta^2}{m+1}(1 - j^2) + \theta^2 j(j+1). \tag{47}$$

(For similar information in this context see, Basak and Balakrishnan (2003)). For all combinations the values of m and n, the MLP of Y_j based on **Plan B** in (28) is more efficient than the MLP of R_j based on **Plan A** if and only if

$$RE = \frac{MSPE(\hat{Y}_j^P(B), Y_j)}{MSPE(\hat{R}_j^P(A), R_j)} < 1. \tag{48}$$

After some simplification, this inequality holds if and only if $\frac{m}{N}(\frac{N+1}{m+1}) < 1$. The values of RE for $j = 8$, $m = 3(1)7$ and $N = 6, 10, 15, 21, 28$ (or equivalently n=3,4,5,6,7) are plotted in Figure 1. From Figure 1, it is easy to observe that for $\frac{m}{N}(\frac{N+1}{m+1}) \geq 1$, we have $RE \geq 1$. This figure shows that for given m when n increases the RE decreases and is smaller than 1.

In Figures 2 and 3, from (12) and (38), we have plotted the values of $RE = \frac{MSPE(\hat{Y}_j^{(BS)},Y_j)}{MSPE(\hat{R}_j^{(AS)},R_j)}$ versus m for given $j = 8$ and some different choices of m and N. First of all, we consider two cases for the prior parameters (α, β): (2,2) and (0.0001,1000), which for ease are indicated by informative and non-informative priors, respectively. From Figure 2, it is observed that, under SE loss function and informative prior, for given m when n increase, the RE

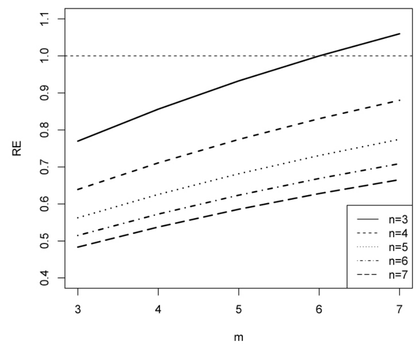

Figure 2. The plot of $RE = \frac{MSPE(\hat{Y}_j^{(BS)}, Y_j)}{MSPE(\hat{R}_j^{(AS)}, R_j)}$ versus m under informative prior for $j = 8$, $m = 3(1)7$ and $N = 6, 10, 15, 21, 28$.

decreases and is smaller than 1, i.e. under SE loss function and informative prior, the Bayes point predictors obtained based on **Plan B** have smaller MSPEs as compared to **Plan A**. But for all combinations the values of m and n and $\alpha = \beta = 2$, the Bayes point predictors based on **Plan B** in (11) is more efficient than the Bayes point predictors based on **Plan A** given by (37) if and only if

$$MSPE\big(\hat{Y}_j^{(BS)}, Y_j\big) < MSPE\big(\hat{R}_j^{(AS)}, R_j\big). \tag{49}$$

This inequality holds if and only if $(\frac{m+1}{N+1})^2 < (\frac{1-4\theta-4N\theta^2-4N^2\theta^2}{1-4\theta-4m\theta^2-4m^2\theta^2})$. Moreover, Figure 3 shows the corresponding result under SE loss function and non-informative prior, which is similar to that of case informative prior (Figure2).

6 Illustrative Example

In order to illustrate the results obtained in the preceding sections, we analyze a real data set. Here, we consider a data set (say Y) on the daily heat degree of the month of January for seven years 2006, 2007, 2013, 2014, 2017, 2018, 2019 of Quebec Province and AMQUI Station (the given data are available at this address: first, log in to the website: www.climate.weather.gc.ca

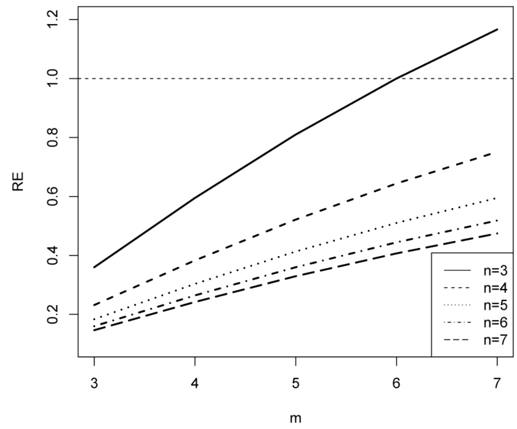

Figure 3. The plot of $RE = \frac{MSPE(\hat{Y}_j^{(BS)}, Y_j)}{MSPE(\hat{R}_j^{(AS)}, R_j)}$ versus m under non-informative prior for $j = 8$, $m = 3(1)7$ and $N = 6, 10, 15, 21, 28$.

and click on the Historical Data. The web site of Historical Data will then appear). The Weibull distribution was one of the best models fitted on Y based on suggestions of easy fit software. In more detail, based on the maximum likelihood approach, we have $Y \sim Weibull(a = 4.76, b = 32.57)$ with pdf $f(y) = \frac{a}{b} y^{\frac{1}{b}-1} e^{-ay^{\frac{1}{b}}}$. Consequently, $X = aY^{\frac{1}{b}} \sim Exp(\theta)$. The p-value of the goodness of Kolmogorov-Smirnov fit test of the one-parameter exponential distribution on the converted data, X, is 0.998, which supports the adequacy of the fitting. The results obtained in the preceding sections can be applied to X. The records extracted from this transformed data set are $\mathbf{r_A} = (5.322, 5.323, 5.357, 5.375)$, and $\mathbf{r_B} = (5.322, 5.375, 5.340, 5.334, 5.350, 5.348, 5.337)$, so we have $m = 4$ and $N = 28$ (or n=7). We also have $T = \sum_{i=1}^{n} R_{i,i} = 37.41$ and $R_m = 5.375$. It is observed that $R_{i,i}$'s are not necessarily ordered (contrary to R_i's), as mentioned earlier in the introduction section. Based on this starting, according to the Plans **A** and **B**, the corresponding results of the maximum likelihood, as well as Bayes point predictors for the jth upper record value from a future sequence under SE and LINEX loss functions, are reported in Table 1. Since the MSPEs and PRs are a function of θ, we estimated the values of the MSPEs and PRs by substituting their respective the Bayes estimator of θ. In **Plan B**, the Bayes estimator of θ under SE and LINEX loss functions are derived as

$$\hat{\theta}^{(BS)} = E(\theta|\mathbf{r_B}) \qquad \& \qquad \hat{\theta}^{(BL)} = -\frac{1}{a}\ln E(e^{-a\theta}|\mathbf{r_B}),$$

respectively. The Bayes estimator of θ under the LINEX loss function does not have an explicit solution. It must be solved by an appropriate numerical method. Clearly, the PR under the LINEX loss function also does not have a closed-form. To this end, to evaluate the performance of predictors, the parametric bootstrap method is employed. We can use the following algorithm to find the PR:
Step 1. For given values of j, n, α, β and $\mathbf{r_B}$, calculate $\hat{Y}_j^{(BL)}$ and $\hat{\theta}_P(B)$ from (13) and (29) respectively.
Step 2. Generate $r_{i,i} \sim Gamma(i, \hat{\theta}_P(B))$, $i = 1, ..., n$.
Step 3. Calculate $T_b \sim Gamma(N, \hat{\theta}_P(B))$ and $\hat{Yb}_j^{(BL)}$.
Step 4. Repeat Step 3 for $b = 1, ..., B$.
Step 5. Calculate PR=$\frac{1}{B}\sum_{b=1}^{B}[e^{a(\hat{Yb}_j^{(BL)}-\hat{Y}_j^{(BL)})} - a(\hat{Yb}_j^{(BL)} - \hat{Y}_j^{(BL)}) - 1]$.

The point predictors and their MSPEs and PRs have displayed in Table 1. Also, the values of the parameter in the LINEX loss function are considered to be $a = 1, 2$. Furthermore, the effect of the parameter j on $RE = \frac{MSPE(\hat{Y}_j^{(BS)}, Y_j)}{MSPE(\hat{R}_j^{(AS)}, R_j)}$ under SE loss function and, informative and non-informative priors for given $N = 6$ and some different choices of m and j are plotted, in Figures 4 and 5, respectively. In Table 2, according to the HPD method all the simulated values of EL and CP with $K = 10000$, are reported when $m = 3(1)7$, $N = 6, 10, 15, 21, 28$ and $j = 8, 9, 10$. Finally, in Figures 6, 7, 8 and 9, the values of $REL = \frac{EL_j(B)}{EL_j(A)}$, $EL_j(A)$ and $EL_j(B)$ by the survival method under informative and non-informative priors are plotted for some m and n. The following points are extracted from Table 1:

- In both plans, all the MSPEs and PRs are increasing with respect to j.
- Moreover, the Bayes predictors under the LINEX loss function performs better than the Bayes predictors under the SE loss function. However, the predictors under the SE loss function performs better than the MLP. So by comparing the results, the superiority of the Bayes predictors in comparison with the classical ones will be revealed.
- According to the obtained results, **Plan B** performs better than **PlanA**.

Table 1. Point predictors for the jth future upper record values and their MSPEs and PRs based on a real data in Plans A and B, when $j = 2, 4, 6, 8, 9, 10$, $m = 4$, $N = 28$.

	Plan A							
	MLP		SEL		LINEX($a = 1$)		LINEX($a = 2$)	
$j \downarrow$	pred	MSPE	pred	MSPE	pred	PR	pred	PR
2	1.07	4.97	2.35	3.72	1.40	0.18	1.09	0.08
4	3.22	11.05	4.70	9.35	2.67	0.62	2.05	0.29
6	5.38	19.33	7.05	16.89	3.82	1.25	2.94	0.59
8	7.52	29.82	9.40	26.35	4.90	2.01	3.76	0.96
9	8.60	35.90	10.57	31.79	5.42	2.44	4.15	1.16
10	9.67	42.52	11.75	37.72	5.92	2.89	4.52	1.38
	Plan B							
	MLP		SEL		LINEX($a = 1$)		LINEX($a = 2$)	
$j \downarrow$	pred	MSPE	pred	MSPE	pred	PR	pred	PR
2	1.29	4.30	2.61	3.65	1.64	0.04	1.26	0.02
4	3.87	9.43	5.23	7.76	3.24	0.17	2.49	0.07
6	6.45	14.02	7.84	12.33	4.80	0.36	3.68	0.15
8	9.03	19.09	10.46	17.36	6.33	0.62	4.85	0.27
9	10.32	21.80	11.76	20.05	7.08	0.78	5.42	0.33
10	11.61	24.63	13.07	22.85	7.83	0.95	5.99	0.41

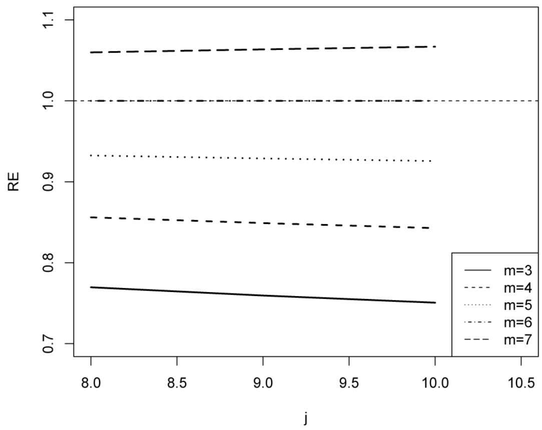

Figure 4. The plot of $RE = \frac{MSPE(\hat{Y}_j^{(BS)}, Y_j)}{MSPE(\hat{R}_j^{(AS)}, R_j)}$ versus j under informative prior for $N = 6$ and some selected of m, j.

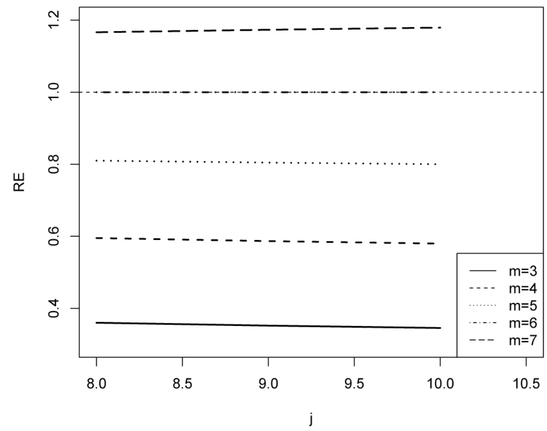

Figure 5. The plot of $RE = \frac{MSPE(\hat{Y}_j^{(BS)}, Y_j)}{MSPE(\hat{R}_j^{(AS)}, R_j)}$ versus j under non-informative prior for $N = 6$ and some selected of m, j.

Figures 4 and 5 show that, respectively, under SE loss function and cases informative and non-informative priors:
• for given N, and $m < 6$ ($m > 6$), when j increase, the RE decreases (increases) slowly and is smaller (higher) than 1. That is for a certain amount of N, and $m < 6$ ($m > 6$), if we increase j, the efficiency of **Plan B** increases (decreases) slowly. Indeed, from Figures 4 and 5, we observe that increasing or decreasing the value of j does not have a significant effect on RE. Also, it is easy to observe that for $m = 6$, we have RE=1.

The simulated values of EL and CP by the HPD method for all combinations of $m, n = 3(1)7$ and $j = 8, 9, 10$ are reported in Table 2. Here, the values in parentheses referes to the CPs. The following points are extracted from Table 2:
• In **PlanB**, for given m, the ELs and CPs improve when n increase, i.e. for given m, the ELs become smaller, and the CPs becom higher, as n increases. Also, we observe that in **PlanA**, the ELs and CPs does not have a special bahaviour.
• In more details, when $m = n$, the expected length of Plan B is uniformly smaller than the expected length of Plan A (the bold values), and also the CP of **PlanB** is more than **PlanA**.

Table 2. The simulated values of EL and CP for the jth future upper record based on **Plans A** and **B** by the HPD method for $m = 3(1)7$, $N = 6, 10, 15, 21, 28$, $j = 8, 9, 10$ and $\gamma = 0.05$.

		Plan A			**Plan B**		
		j			j		
$m\downarrow$	$n\downarrow$	8	9	10	8	9	10
3	3	**16.28**	**17.87**	**19.51**	**16.11**	**17.89**	**19.46**
		(0.814)	(0.806)	(0.808)	(0.896)	(0.891)	(0.899)
	4	16.27	17.82	19.62	15.92	17.31	18.81
		(0.814)	(0.812)	(0.805)	(0.952)	(0.940)	(0.938)
	5	16.03	17.87	19.64	15.38	16.57	17.94
		(0.810)	(0.812)	(0.808)	(0.975)	(0.971)	(0.968)
	6	16.27	17.97	19.60	14.94	16.14	17.29
		(0.815)	(0.812)	(0.806)	(0.987)	(0.987)	(0.982)
	7	16.20	17.70	19.59	14.64	15.83	17.03
		(0.811)	(0.804)	(0.807)	(0.996)	(0.994)	(0.992)
4	3	16.38	17.70	19.85	16.30	17.92	19.56
		(0.852)	(0.849)	(0.845)	(0.900)	(0.901)	(0.892)
	4	**16.23**	**18.03**	**19.84**	**15.85**	**17.28**	**18.83**
		(0.855)	(0.846)	(0.845)	(0.949)	(0.946)	(0.940)
	5	16.18	18.02	19.71	15.28	16.64	17.96
		(0.850)	(0.845)	(0.838)	(0.974)	(0.970)	(0.965)
	6	16.54	18.09	19.73	14.90	16.30	17.48
		(0.856)	(0.848)	(0.851)	(0.987)	(0.987)	(0.983)
	7	16.31	18.08	19.72	14.71	15.82	17.05
		(0.849)	(0.845)	(0.843)	(0.995)	(0.994)	(0.993)
5	3	16.40	18.06	19.60	16.34	17.89	19.47
		(0.881)	(0.870)	(0.871)	(0.905)	(0.893)	(0.896)
	4	16.23	17.83	19.67	15.86	17.27	18.67
		(0.880)	(0.875)	(0.875)	(0.947)	(0.945)	(0.942)
	5	**16.13**	**17.94**	**19.53**	**15.32**	**16.21**	**18.03**
		(0.875)	(0.878)	(0.873)	(0.975)	(0.973)	(0.968)
	6	16.43	17.95	19.19	14.95	16.19	17.38
		(0.880)	(0.878)	(0.863)	(0.988)	(0.986)	(0.983)
	7	16.25	18.14	19.63	14.71	15.85	17.01
		(0.877)	(0.881)	(0.872)	(0.995)	(0.993)	(0.992)
6	3	16.23	17.84	19.60	16.40	17.87	19.58
		(0.904)	(0.896)	(0.895)	(0.907)	(0.897)	(0.894)
	4	16.21	17.95	19.42	15.89	17.39	18.79
		(0.903)	(0.897)	(0.890)	(0.951)	(0.944)	(0.939)
	5	16.35	17.99	19.67	15.35	16.54	17.95
		(0.903)	(0.895)	(0.891)	(0.975)	(0.979)	(0.970)
	6	**16.23**	**17.80**	**19.54**	**15**	**16.15**	**17.43**
		(0.901)	(0.898)	(0.896)	(0.988)	(0.986)	(0.984)
	7	16.31	17.92	19.53	14.69	15.82	17.02
		(0.905)	(0.898)	(0.896)	(0.995)	(0.993)	(0.993)
7	3	16.22	17.76	19.27	16.20	17.72	19.60
		(0.916)	(0.914)	(0.909)	(0.920)	(0.924)	(0.926)
	4	16.14	17.83	19.23	15.84	17.32	18.80
		(0.914)	(0.910)	(0.905)	(0.949)	(0.945)	(0.944)
	5	16.31	17.73	19.40	15.30	16.61	17.88
		(0.915)	(0.914)	(0.914)	(0.973)	(0.970)	(0.967)
	6	16.19	17.84	19.37	14.89	16.19	17.38
		(0.915)	(0.907)	(0.906)	(0.989)	(0.987)	(0.984)
	7	**16.21**	**17.73**	**19.44**	**14.67**	**15.90**	**17.00**
		(0.916)	(0.913)	(0.908)	(0.993)	(0.994)	(0.993)

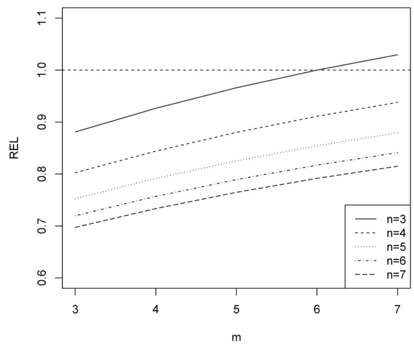

Figure 6. The plot of $REL = \frac{EL_j(B)}{EL_j(A)}$ versus m under informative prior for $j = 8$, $\gamma = 0.05$ and some selected of m and n.

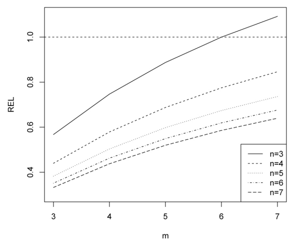

Figure 7. The plot of $REL = \frac{EL_j(B)}{EL_j(A)}$ versus m under non-informative prior for $j = 8$, $\gamma = 0.05$ and some selected of m and n.

• The Bayesian prediction intervals based on **PlanB** as compared to the Bayesian prediction intervals based **PlanA** acts very good, in terms of ELs and CPs i.e., the ELs on the basis of **PlanB** are smaller than those of **PlanA** and the CPs of the **PlanB** is more than **PlanA**.

The following points are extracted from Figures 6 and 7:
• for given m the results shows that under informative prior when n increase, the RE decreases and is smaller than 1.
•In Figure 7, under non-informative prior, we observe that the behavior of $REL = \frac{EL_j(B)}{EL_j(A)}$ is similar to that of case informative prior. The values

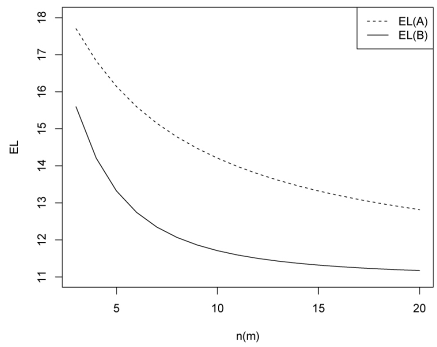

Figure 8. The plot of $EL_j(A)$ and $EL_j(B)$ versus $m = n = 3(1)20$ under informative prior and survival method for $j = 8$ and $\gamma = 0.05$.

of $EL_j(B)$ and $EL_j(A)$ under informative prior and survival method, are plotted in Figure 8 for $j = 8$, $\gamma = 0.05$ and $m = n = 3(1)20$. This figure shows that, when $m = n$, under survival method the expected length based on **PlanB** is uniformly smaller than the expected length based on **PlanA** or equivalently, we have $REL < 1$. When $m = n$, the obtained results in Figure 8 confirm the results of the HPD method (in Table 2, see the bold values). In Figure 9, under non-informative prior, we observe that the behavior of $EL_j(B)$ and $EL_j(A)$ is similar to that of informative prior.

7 Conclusion

In this study, we have considered the behavior of predictors of future record values when the only observed data on the exponential model are upper RRSS's. We compared some of our results with the inverse sampling scheme. To this end, we predicted upper record values from a future sequence using maximum likelihood predictor and Bayes point predictors under squared error (SE) and linear-exponential (LINEX) loss functions. As regard point prediction, results showed that the Bayes predictors under the LINEX loss function performs better than the Bayes predictors under the SE loss function. However, the predictors under the SE loss function performs better than the MLP. So by comparing the results, the superiority of the Bayes

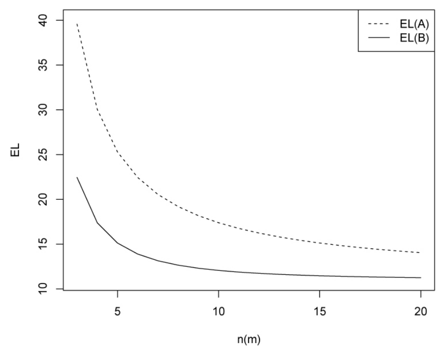

Figure 9. The plot of $EL_j(A)$ and $EL_j(B)$ versus $m = n = 3(1)20$ under non-informative prior and survival method for $j = 8$ and $\gamma = 0.05$.

predictors in comparison with the classical ones will be revealed. Also, it is observed that most of the point predictors based on **PlanB** are superior than **PlanA**. Next, two Bayesian prediction intervals for future upper record values are derived. As regard Bayesian prediction intervals, we observed that, almost everywhere under the survival method, the Bayesian prediction intervals perform better in **PlanB**. Furthermore, under the HPD method, Bayesian prediction intervals obtained based on **PlanB** has a smaller expected length as compared to **PlanA**. Also, Bayesian prediction intervals obtained based on **PlanB** has further coverage probability as compared to Plan A. According to the obtained results of the survival and HPD methods, it is observed that, for $m = n$, the expected length based on **PlanB** is uniformly smaller than the expected length based on **PlanA**. Numerical results showed that in most of the situations that **Plan B** provides more efficient predictors than their counterparts. It is recommended to use the RRSS scheme if conditions are ready for the RRSS scheme.

References

Ahmadi, J., MirMostafaee, S.M.T.K. and Balakrishna, N. (2012). Bayesian Prediction of k-record Values Based on Progressively Censored Data from Exponential Distribution. *Journal of Statistical Computation and Simulation*, **82**, 51-62.

Ahsanullah, M. (1995). *Record Statistics.* Nova Science Publishers Inc, Commack, New York.

Aitchison, J. and Dunsmore, I.R. (1975). *Statistical Prediction Analysis.* Cambridge University Press, Cambridge.

Arnold, B.C., Balakrishnan, N. and H. Nagaraja, N. (1998). *Records.* John Wiley & Sons, New York.

Basak, P. and Balakrishnan, N. (2003). Maximum Likelihood Prediction of Future Record Statistic, In *Mathematical and Statistical Method in Reliability, Series on Quality, Reliability and Engineering*, Volume **7**, 159-175.

Casella, G. and Berger, R.L. (2002). *Statistical Inference.* 2nd ed., Duxbury Press, the University of Michigan.

Chen, M. and Shao, Q. (1999). Monte Carlo Estimation of Bayesian Credible and HPD Intervals. *Comput Graph Stat*, **8**, 69-92.

Eskandarzadeh, M., Tahmasebi, S. and Afshari, M. (2016). Information Measures for Record Ranked Set Samples. *Ciencia e Natura*, **38**, 554-563.

Geisser, S. (1993). *Predictive Inference: An Introduction.* Chapman and Hall.

Gulati, S., and Padgett, W.J. (2003). *Parametric and Non-parametric Inference from Record-breaking Data.* Lecture Notes In Statistics, No. 172, Springer, New York.

Kaminsky, K.S., Rhodin, L.S. (1985). Maximum Likelihood Prediction. *Ann. Inst. Statist. Math.*, **37**, 507-517.

McIntyer GA.A. (1952). Method for Unbiased Selective Sampling, Using Ranked Sets. *Aust. J. Agric. Res.*, **3**, 385-390.

MirMostafaee, S.M.T.K. and Ahmadi, J. (2011). Point Prediction of Future Order Statistics from an Exponential Distribution. *Statistics and Probability Letters*, **81**, 360-370.

Nevzorov, V.B. (2001).*Records.* Mathematical Theory, Translation of Mathematical Monographs, American Mathematical Society, Providence Rhode Island.

Paul, J. and P.Y. Thomas. (2017). Concomitant Record Ranked Set Sampling. *Communications in Statistics-Theory and Methods*, **46**, 9518-9540.

Raqab, M. and Balakrishnan, N. (2008). Prediction Interval for Future Records. *Statistics and Probability,* **78**, 1955-1963.

Sadeghpour, A., Salehi, M. and Nezakati, A. (2020). Estimations of the Stress-strength Reliability Using Lower Record Ranked Set Sampling Scheme Under the Generalized Exponential Distribution. *Statistical Computation and Simulation*, **90**, 51-74.

Safaryian, A., Arashi, M. and Arabi Belaghi R. (2019). Improved Estimators for Stress-strength Reliability Using Record Ranked Set Sampling Scheme. *Communications in Statistics-Simulation and Computation,* **48**, 2708-2726.

Salehi, M. and Ahmadi, J. (2014). Record Ranked Set Sampling Scheme. *Metron*, **72**, 351-365.

Salehi, M. and Ahmadi, J. (2015). Estimation of Stress-strength Reliability Using Record Ranked Set Sampling Scheme from the Exponential Distribution. *Filomat*, **29**, 1149-1162.

Salehi, M., Ahmadi, J. and Balakrishnan, N. (2015). Prediction of Order Statistics and Record Values based on Ordered Ranked Set Sampling. *Statistical Computation and Simulation*, **85**, 77-88.

Salehi, M., Ahmadi, J. and Dey, S. (2016). Comparison of Two Sampling Schemes for Generating Record-breaking Data from the Proportional Hazard Rate Models. *Communications in Statistics-Theory and Methods*, **45**, 3721-3733.

Varian, H.R. (1975). *A Bayesian Approach to Real Estate Assessment.* Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. North Holland, Amsterdam, 195-208.

Zellner, A. (1986). Bayesian Estimation and Prediction Using Asymmetric Loss Function. *J. Am. Stat. Assoc.*, **81**, 446-451.

Ehsan Golzade Gervi
Department of Statistics,
University of Payame Noor,
Tehran, Iran.
email: *e_golzade_g@yahoo.com*

Parviz Nasiri
Department of Statistics,
University of Payame Noor,
Tehran, Iran.
email: *pnasiri45@yahoo.com*

Mahdi Salehi
Department of Mathematics and Statistics,
University of Neyshabur,
Neyshabur Iran.
email: *salehi.sms@neyshabur.ac.ir*