One-sequence and two-sequence prediction for future Weibull records

Omar M. Bdair*

Faculty of Engineering Technology, Al-Balqa Applied University
Amman 11134, Jordan
bdairmb@yahoo.com

Mohammad Z. Raqab

Department of Statistics and OR, Kuwait University
Safat 13060, Kuwait
Department of Mathematics, University of Jordan
Amman 11942, Jordan
mraqab@ju.edu.jo

Based on record data, prediction of the future records from the two-parameter Weibull distribution is studied. First we consider the sampling based procedure to compute the Bayes estimates and also to construct symmetric credible intervals. Secondly, we consider one-sequence and two-sequence Bayes prediction of the future records based on some observed records. The Monte Carlo algorithms are used to compute simulation consistent predictors and prediction intervals for future unobserved records. A numerical simulation study is conducted to compare the different methods and a real data set involving the annual rainfall recorded at Los Angeles Civic Center during 132 years is analyzed to illustrate the procedures developed here.

Keywords: Weibull distribution; record values; Bayes estimation; Bayes prediction; Monte Carlo samples.

2000 Mathematics Subject Classification: 62G30, 62F15, 62F10.

1. Introduction

The Weibull distribution is one of the most popular distributions used for analyzing skewed lifetime data. A detailed discussion on it has been provided by ([16], Chapter 21). Because of the various shapes of the PDF and due to monotone (increasing/decreasing) property of the hazard function, it has been used as an alternative to the gamma or log-normal distribution. The Weibull distribution with the shape and scale parameters α and λ will be denoted by $WE(\alpha, \lambda)$. Let $X_1, X_2, ...$ be a sequence of independent and identically distributed (iid) random variables from two-parameter WE

---

*Faculty of Engineering Technology, Al-Balqa Applied University, Amman 11134, Jordan.

distribution. The two-parameter WE distribution has CDF

$$F_{WE}(x; \alpha, \lambda) = 1 - e^{-\lambda x^\alpha}, \text{ for } x > 0, \tag{1.1}$$

and the PDF

$$f_{WE}(x; \alpha, \lambda) = \alpha \lambda x^{\alpha-1} e^{-\lambda x^\alpha}, \text{ for } x > 0. \tag{1.2}$$

Here $\alpha > 0, \lambda > 0$ are the shape and scale parameters, respectively.

Record statistics arise naturally in many practical problems, and there are several situations pertaining to meteorology, hydrology, sporting and athletic events wherein only record values may be recorded. For an elaborate treatment on records and their applications, one may refer to the books by [6], [4] and [17]. For a detailed discussion on the Weibull records and associated inferences, one may refer to [22].

Prediction problems have received considerable attention in the last decades. Several real applications can be found in actuarial studies, warranty data analysis, rainfall extremes and highest water levels. For example, when setting up warranty period for a product, a manufacturer would use some of the known previous failure times to predict a suitable warranty period of the product. Extensive work on prediction can be found in the literature. [13] used conditional argument to develop prediction intervals for future Weibull order statistics. [5] also considered the Bayesian prediction problem for a large class of lifetime distributions. A numerical approach to Bayesian prediction for the two-parameter WE distribution was considered by [10]. They provided numerical Bayes predictor of the future observation, under the assumptions that the shape parameter has a uniform prior over a finite interval and the error is squared error. [14] presented Bayes prediction intervals for future observations are obtained in the one-and two-sample cases for Weibull model. Classical and Bayesian prediction of future observations based on hybrid censored samples from Weibull distribution are considered by [8].

In the context of record data, [3] considered the problem of predicting the future records based on some observed ones when the observations are from the exponential distribution. [15] developed prediction bounds for future generalized exponential lower records by using Bayes and empirical Bayes techniques. [20] discussed the Bayesian analysis in the context of records from the two-parameter WE distribution through using a discrete prior for the shape parameter. Bayesian estimation and prediction for some life distributions based on record values, including Exponential, Weibull, Pareto and Burr type XII, are discussed in [2]. [21] obtained the best linear unbiased predictors of the next lower records based on observing the first n lower records from gamma distribution. [7] considered prediction of future record values from the exponentiated family of distributions from a Bayesian view point.

In the Bayesian analysis, the performance of the predictor depends on the prior distribution and also on the loss function chosen. For example, [19] used a continuous-discrete joint prior distribution. Another prior is a conjugate prior on λ and independent uniform prior on α. Some authors used independent uniform priors on both α and λ (see [10]) in the respect. For estimating θ by a decision δ, the following loss functions can be considered. The first is the squared loss function

defined as

$$L_1(\theta, \delta) = (\theta - \delta)^2.$$

The squared error loss is a symmetric function that penalizes overestimation and underestimation equally, and takes the value zero when the estimate is right on target.

There are many alternatives to the above loss function. Among them, the absolute loss function which is defined by

$$L_2(\theta, \delta) = |\theta - \delta|.$$

This loss function may be considered as a reasonable alternative to the squared loss function, but mathematically analysis based on absolute error is often substantially less tractable than that based on squared error. Both loss functions capture the basic idea that one is penalized, in a symmetric fashion, for the distance between an estimator and its target.

Other alternatives are to consider versions of asymmetric loss functions. One of the more widely used forms of asymmetric loss is the LINEX (or, "linear-exponential") loss function, the most common version of which is the one proposed by [23] and defined by

$$L_3(\theta, \delta) = \left(\frac{\delta}{\theta}\right)^{a^*} - a^* \ln\left(\frac{\delta}{\theta}\right) - 1, \, a^* \neq 0.$$

where a^* is the shape parameter of the loss function. It controls the direction and degree of symmetry. When $a^* > 0$, the LINEX loss function imposes a substantial penalty for overestimation. The mirror image of this symmetry may be achieved by setting $a^* < 0$. Another loss function is called entropy loss function (L_4). It is a special case of the LINEX loss function and can be obtained by substituting $a^* = 1$ [see [18]].

In order to perform a Bayesian estimation of the parameters α and λ of Weibull distribution, the prior distribution of α and λ must be specified either when the shape parameter α known or unknown. When the shape parameter α is known, the prior on λ is the conjugate gamma prior. The prior distribution Gamma(a,b) of the scale parameter λ is given by

$$\pi_1(\lambda | a, b) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda}, \, \lambda > 0. \tag{1.3}$$

Here the hyper-parameters $a > 0$, $b > 0$. When both parameters are unknown, following the approach of [9], it is assumed that λ has the same prior (1.3), the prior on α, $\pi_2(\cdot)$ is independent of $\pi_1(\lambda)$, and the PDF of $\pi_2(\alpha)$ is log-concave on the support $(0, \infty)$.

Our aim in this paper is to predict the future records based on some observed WE records. Here, we consider the one-sequence and two-sequence prediction problems for future WE records. The sample-based procedure is adopted in this current work to estimate the parameters and then estimate the predictive density of future WE records. The main additional difference between our work and the previous Bayesian works mentioned above is that the proposed priors are general ones. Further, we have chosen different error loss functions. In this context, we have used a Monte Carlo (MC) samples to compute the predictive density and then to compute the Bayes predictors (BPs) as well as the corresponding prediction intervals (PIs).

The rest of the paper is organized as follows. In Section 2, we provide sample-based estimates for the predictive density functions of the future records based on one-sequence prediction problem.

The two-sequence prediction problem is discussed in Section 3. Section 4 presents a data analysis to illustrate our results, and a MC simulation that compares the BPs under informative with those under non-informative priors. Finally, we conclude our work in Section 5.

2. One-sequence prediction problem

In this section, we consider the BPs of future records based on observed records under different loss functions. Here, the future record to be predicted comes from the same sequence of observed records and it is therefore correlated with the observed record data. Suppose that we observe only the first m upper records $\underset{\sim}{x} = (x_{U(1)}, x_{U(2)}, ..., x_{U(m)})$. Here, we obtain the BPs as well as constructing the PIs for the nth future upper record $X_{U(n)}$ under different loss functions, where $1 \leq m < n$. The posterior predictive density of $X_{U(n)}$ can be written as

$$\pi_{X_{U(n)}}(y|\underset{\sim}{x}) = \int_0^\infty \int_0^\infty f_{X_{U(n)}}(y|\underset{\sim}{x}, \alpha, \lambda) \, \pi(\alpha, \lambda|\underset{\sim}{x}) \, d\alpha \, d\lambda, \; y > x_{U(m)},$$

where $f_{X_{U(n)}}(y|\underset{\sim}{x}, \alpha, \lambda)$ is the conditional density function of $X_{U(n)}$ given the records data $\underset{\sim}{x}$. Using the fact that the record values satisfy Markov property, the conditional PDF of $X_{U(n)}$ given $\underset{\sim}{x}$ is just the conditional PDF of $X_{U(n)}$ given $x_{U(m)}$, *i.e.*

$$f_{X_{U(n)}}(y|\underset{\sim}{x}, \alpha, \lambda) = f_{X_{U(n)}}(y|x_{U(m)}, \alpha, \lambda) = \frac{[H(x_{U(n)}) - H(x_{U(m)})]^{n-m-1}}{(n-m-1)!} \frac{f(x_{U(n)})}{1 - F(x_{U(m)})},$$
$$-\infty < x_{U(m)} < x_{U(n)} < \infty,$$

where $H(x) = -\ln(1 - F(x))$. Based on the PDF and CDF of Weibull distribution and the binomial expansion, the conditional PDF of $X_{U(n)}$ given $X_{U(m)} = x_{U(m)}$ can be written as follows

$$f_{X_{U(n)}}(y|x_{U(m)}, \alpha, \lambda) = \frac{\alpha \lambda^{n-m} e^{\lambda x_{U(m)}^\alpha}}{(n-m-1)!} \sum_{i=0}^{n-m-1} \binom{n-m-1}{i} (-1)^{n-m-i-1} x_{U(m)}^{(n-m-i-1)\alpha}$$
$$\times y^{\alpha(i+1)-1} e^{-\lambda y^\alpha}, \; y > x_{U(m)}. \tag{2.1}$$

The posterior predictive density of $X_{U(n)}$ at any point $y > x_{U(m)}$ is then

$$\pi_{X_{U(n)}}(y|\underset{\sim}{x}) = E_{posterior}\left[f_{X_{U(n)}}(y|x_{U(m)}, \alpha, \lambda) \right]$$
$$= \int_0^\infty \int_0^\infty \frac{\alpha \lambda^{n-m} e^{\lambda x_{U(m)}^\alpha}}{(n-m-1)!} \sum_{i=0}^{n-m-1} \binom{n-m-1}{i} (-1)^{n-m-i-1} x_{U(m)}^{(n-m-i-1)\alpha}$$
$$\times y^{\alpha(i+1)-1} e^{-\lambda y^\alpha} \, \pi(\alpha, \lambda|\underset{\sim}{x}) \, d\alpha d\lambda, \; y > x_{U(m)}. \tag{2.2}$$

Under the squared error loss function L_1, the BP of $Y = X_{U(n)}$ can be obtained as

$$
\begin{aligned}
X_{U(n)}^{BP1} &= E(Y|\underset{\sim}{x}) \\
&= \int\limits_{x_{U(m)}}^{\infty} \left[\int\limits_0^{\infty} \int\limits_0^{\infty} \frac{\alpha \lambda^{n-m} e^{\lambda x_{U(m)}^{\alpha}}}{(n-m-1)!} \sum_{i=0}^{n-m-1} \binom{n-m-1}{i} (-1)^{n-m-i-1} x_{U(m)}^{(n-m-i-1)\alpha} \right. \\
&\quad \left. \times\, y^{\alpha(i+1)} e^{-\lambda y^{\alpha}} \pi(\alpha, \lambda|\underset{\sim}{x}) \, d\alpha d\lambda \right] dy.
\end{aligned}
$$

By using MC samples $\{(\alpha_j, \lambda_j); \ j = 1, 2, ..., M\}$ obtained using the algorithm in Appendix C and the PDF of the gamma distribution, a simulation consistent based predictor $\hat{X}_{U(n)}^{BP1}$ of $Y = X_{U(n)}$ can be given as

$$
\begin{aligned}
\hat{X}_{U(n)}^{BP1} &= \frac{1}{M} \sum_{j=1}^{M} \frac{\lambda_j^{n-m-1-\frac{1}{\alpha_j}} e^{\lambda_j x_{U(m)}^{\alpha_j}}}{(n-m-1)!} \sum_{i=0}^{n-m-1} \binom{n-m-1}{i} (-1)^{n-m-i-1} x_{U(m)}^{(n-m-i-1)\alpha_j} \\
&\quad \times \frac{\gamma\left(\frac{1}{\alpha_j}+i+1, \ \lambda_j x_{U(m)}^{\alpha_j}\right)}{\lambda_j^i}.
\end{aligned}
\tag{2.3}
$$

The simulation consistent based predictor of the first unobserved record value can be obtained by setting $n = m + 1$ in Eq.(2.3) as follows:

$$
\hat{X}_{U(m+1)}^{BP1} = \frac{1}{M} \sum_{j=1}^{M} \lambda_j^{-\frac{1}{\alpha_j}} e^{\lambda_j x_{U(m)}^{\alpha_j}} \gamma\left(\frac{1}{\alpha_j}+1, \ \lambda_j x_{U(m)}^{\alpha_j}\right).
$$

Under the absolute error loss function L_2, the corresponding BP of $X_{U(n)}$, $1 \le m < n$, denoted by $X_{U(n)}^{BP2}$ is the median of the posterior predictive density of $X_{U(n)}$, Eq.(2.2), which is obtained by solving the following equation with respect to $X_{U(n)}^{BP2}$

$$
\int\limits_{x_{U(m)}}^{X_{U(n)}^{BP2}} \pi_{X_{U(n)}}(y|\underset{\sim}{x}) \, dy = \frac{1}{2} \Leftrightarrow \int\limits_{X_{U(n)}^{BP2}}^{\infty} \pi_{X_{U(n)}}(y|\underset{\sim}{x}) \, dy = \frac{1}{2}.
\tag{2.4}
$$

Based on the posterior predictive density of $X_{U(n)}$, Eq.(2.2), Eq.(2.4) is equivalent to

$$
\begin{aligned}
\int\limits_{X_{U(n)}^{BP2}}^{\infty} \int\limits_0^{\infty} \int\limits_0^{\infty} &\left[\frac{\alpha \lambda^{n-m}}{(n-m-1)!} \sum_{i=0}^{n-m-1} \binom{n-m-1}{i} (-1)^{n-m-i-1} x_{U(m)}^{(n-m-i-1)\alpha} e^{\lambda x_{U(m)}^{\alpha}} y^{\alpha(i+1)-1} e^{-\lambda y^{\alpha}} \right] \\
&\times \pi(\alpha, \lambda|\underset{\sim}{x}) \, d\alpha d\lambda \, dy = \frac{1}{2}.
\end{aligned}
$$

Based on MC samples $\{(\alpha_j, \lambda_j); \ j = 1, 2, ..., M\}$, a simulation consistent based predictor $\hat{X}_{U(n)}^{BP_2}$ of $Y = X_{U(n)}$ can be found by solving the following nonlinear equation with respect to $\hat{X}_{U(n)}^{BP_2}$

$$\frac{1}{M} \sum_{j=1}^{M} \frac{e^{\lambda_j x_{U(m)}^{\alpha_j}}}{(n-m-1)!} \sum_{i=0}^{n-m-1} \binom{n-m-1}{i} (-\lambda_j)^{n-m-i-1} x_{U(m)}^{(n-m-i-1)\alpha_j}$$

$$\times \gamma \left(i+1, \left(\lambda_j (\hat{X}_{U(n)}^{BP2})^{\alpha_j} \right) \right) = \frac{1}{2}. \qquad (2.5)$$

Under the LINEX loss function L_3, the corresponding simulation BP $X_{U(n)}^{BP_3}$ of $Y = X_{U(n)}$ can be obtained as

$$X_{U(n)}^{BP_3} = \left[E(Y^{-a^*} | \underset{\sim}{x}) \right]^{-\frac{1}{a^*}}$$

$$= \left[\int_{x_{U(m)}}^{\infty} \int_0^{\infty} \int_0^{\infty} \frac{\alpha \lambda^{n-m}}{(n-m-1)!} \sum_{i=0}^{n-m-1} \binom{n-m-1}{i} (-1)^{n-m-i-1} x_{U(m)}^{(n-m-i-1)\alpha} \right.$$

$$\left. \times y^{\alpha(i+1)-a^*-1} e^{-\lambda(y^\alpha - x_{U(m)}^\alpha)} \pi(\alpha, \lambda | \underset{\sim}{x}) \, d\alpha d\lambda \, dy \right]^{-\frac{1}{a^*}}.$$

Based on MC samples $\{(\alpha_j, \lambda_j); \ j = 1, 2, ..., M\}$ and the PDF of the gamma distribution, a simulation consistent based predictor $\hat{X}_{U(n)}^{BP_3}$ of $X_{U(n)}$ will be

$$\hat{X}_{U(n)}^{BP_3} = \left[\frac{1}{M} \sum_{j=1}^{M} \frac{\lambda_j^{n-m-1+\frac{a^*}{\alpha_j}} e^{\lambda_j x_{U(m)}^{\alpha_j}}}{(n-m-1)!} \sum_{i=0}^{n-m-1} \binom{n-m-1}{i} (-1)^{n-m-i-1} x_{U(m)}^{(n-m-i-1)\alpha_j} \right.$$

$$\left. \times \frac{\gamma \left(i - \frac{a^*}{\alpha_j} + 1, \lambda_j x_{U(m)}^{\alpha_j} \right)}{\lambda^i} \right]^{-\frac{1}{a^*}}. \qquad (2.6)$$

In particular, the simulation BP $\hat{X}_{U(m+1)}^{BP3}$ of $X_{U(m+1)}$ can be obtained by setting $n = m+1$ in Eq.(2.6) as follows :

$$\hat{X}_{U(m+1)}^{BP3} = \left[\frac{1}{M} \sum_{j=1}^{M} \lambda_j^{\frac{a^*}{\alpha_j}} e^{\lambda_j x_{U(m)}^{\alpha_j}} \gamma(1 - \frac{a^*}{\alpha_j}, \lambda_j x_{U(m)}^{\alpha_j}) \right]^{-\frac{1}{a^*}}.$$

To obtain prediction bounds on $Y = X_{U(n)}$, $1 \le m < n$, under different loss functions, we need to find the predictive survival function of $Y = X_{U(n)}$ at any point $y > x_{U(m)}$. The predictive survival function is defined by

$$S_{X_{U(n)}}(y | \underset{\sim}{x}) = E_{posterior} \left(S_{X_{U(n)}}(y | \underset{\sim}{x}, \alpha, \lambda) \right)$$

$$= \int_0^{\infty} \int_0^{\infty} S_{X_{U(n)}}(y | \underset{\sim}{x}, \alpha, \lambda) \pi(\alpha, \lambda | \underset{\sim}{x}) \, d\alpha d\lambda, \qquad (2.7)$$

where $S_{X_{U(n)}}(y|\underset{\sim}{x}, \alpha, \lambda)$ is the survival function of $Y = X_{U(n)}$ given $\underset{\sim}{x}$. Using the Markovian property of the records, we have

$$S_{X_{U(n)}}(y|\underset{\sim}{x}, \alpha, \lambda) = P(Y > y|x_{U(m)}, \alpha, \lambda)$$

$$= \int_y^\infty \frac{[H(z) - H(x_{U(m)})]^{n-m-1}}{(n-m-1)!} \frac{f(z|\alpha, \lambda)}{1 - F(x_{U(m)}|\alpha, \lambda)} dz.$$

By making the transformation $v = H(z) - H(x_{U(m)})$ and using the relation between the incomplete gamma function and sum of poisson probabilities, we have

$$S_{X_{U(n)}}(y|x_{U(m)}, \alpha, \lambda) = \sum_{j=0}^{n-m-1} \frac{e^{-\lambda(y^\alpha - x_{U(m)}^\alpha)} [\lambda(y^\alpha - x_{U(m)}^\alpha)]^j}{j!}. \tag{2.8}$$

The predictive survival function for $Y = X_{U(n)}$, Eq.(2.7), becomes

$$S_{X_{U(n)}}(y|\underset{\sim}{x}) = \int_0^\infty \int_0^\infty \left[\sum_{j=0}^{n-m-1} \frac{e^{-\lambda(y^\alpha - x_{U(m)}^\alpha)} [\lambda(y^\alpha - x_{U(m)}^\alpha)]^j}{j!} \right] \pi(\alpha, \lambda|\underset{\sim}{x}) d\alpha d\lambda. \tag{2.9}$$

Notice that Eq.(2.9) can't be expressed in a closed form and hence can't be evaluated analytically. By using the MC samples $\{(\alpha_i, \lambda_i); i = 1, 2, ..., M\}$ and under the square error loss function L_1, the simulation consistent based predictor of the predictive survival function for $X_{U(n)}$ will be

$$\hat{S}_{X_{U(n)}}(y|\underset{\sim}{x}) = \frac{1}{M} \sum_{i=1}^M \left[\sum_{j=0}^{n-m-1} \frac{e^{-\lambda_i(y^{\alpha_i} - x_{U(m)}^{\alpha_i})} [\lambda_i(y^{\alpha_i} - x_{U(m)}^{\alpha_i})]^j}{j!} \right].$$

Under the absolute error loss function L_2, the simulation consistent based predictor of the predictive survival function for $X_{U(n)}$, can be obtained by using the following algorithm:

Algorithm 1

- **Step 1:** Evaluate $S = S_{X_{U(n)}}(y|x_{U(m)}, \alpha, \lambda)$, Eq.(2.8), at each sample (α_i, λ_i), for $i = 1, 2, ..., M$, to get $S_1, S_2, ..., S_M$

- **Step 2:** Order $S_1, S_2, ..., S_M$ as $S_{(1)} < S_{(2)} < ... < S_{(M)}$.

- **Step 3:** The simulation consistent based predictor of the predictive survival function for $X_{U(n)}$ is given by

$$\hat{S}_{X_{U(n)}}(y|\underset{\sim}{x}) = Median\left[S_{(1)}, S_{(2)}, ..., S_{(M)}\right]$$

Under the LINEX loss function L_3, the simulation consistent based predictor of the predictive survival function for $X_{U(n)}$ can be obtained as

$$\hat{S}_{X_{U(n)}}(y|\underset{\sim}{x}) = \left[\frac{1}{M} \sum_{i=1}^{M} \left(\sum_{j=0}^{n-m-1} \frac{e^{-\lambda_i(y^{\alpha_i} - x_{U(m)}^{\alpha_i})} [\lambda_i(y^{\alpha_i} - x_{U(m)}^{\alpha_i})]^j}{j!} \right)^{-a^*} \right]^{-\frac{1}{a^*}}.$$

A $(1-\beta)100\%$ PI for $X_{U(n)}$, $1 \leq m < n$, can be obtained by solving the non-linear equations (2.10) and (2.11) for the lower bound L and upper bound U :

$$P(X_{U(n)} > L|\underset{\sim}{x}) = 1 - \frac{\beta}{2} \iff \hat{S}^P_{X_{U(n)}}(L|\underset{\sim}{x}) = 1 - \frac{\beta}{2}, \tag{2.10}$$

$$P(X_{U(n)} > U|\underset{\sim}{x}) = \frac{\beta}{2} \iff \hat{S}^P_{X_{U(n)}}(U|\underset{\sim}{x}) = \frac{\beta}{2}. \tag{2.11}$$

Since Eqs. (2.10) and (2.11) can't be solved explicitly, a suitable numerical method should be used.

3. Two-sequence prediction problem

Let $\underset{\sim}{x} = (x_{U(1)}, x_{U(2)}, \ldots, x_{U(m)})$ be the first m observed records from a sequence from WE(α, λ). Let $y_{U(1)}, y_{U(2)}, \ldots, y_{U(n)}$ be the first n record values from another independent sequence from the same distribution. Based on the observed record sequence, we are interested in predicting the *kth* upper record value $Y_{U(k)}$, $1 \leq k \leq n$, of the future sequence, and obtaining PI of $Y_{U(k)}$. The PDF of the *kth* upper record value $Y_{U(k)}$ [[6] and [4]] is given by

$$g_{(k)}(y|\alpha, \lambda) = \frac{\alpha \lambda^k}{\Gamma(k)} y^{\alpha k - 1} e^{-\lambda y^\alpha}. \tag{3.1}$$

To obtain the Bayes predictive estimator of $Y = Y_{U(k)}$, $1 \leq k \leq n$, under different loss functions, we need the posterior predictive density of $Y_{U(k)}$. The posterior predictive density of $Y_{U(k)}$ is denoted by $\pi_{Y_{U(k)}}(y|\underset{\sim}{x})$ and given by

$$\pi_{Y_{U(k)}}(y|\underset{\sim}{x}) = E_{posterior}\left[g_{(k)}(y|\alpha, \lambda) \right]$$
$$= \int_0^\infty \int_0^\infty \frac{\alpha \lambda^k}{\Gamma(k)} y^{\alpha k - 1} e^{-\lambda y^\alpha} \pi(\alpha, \lambda|data) \, d\alpha \, d\lambda. \tag{3.2}$$

Under the square error loss function L_1, the Bayes predictive estimator of $Y = Y_{U(k)}$ can be obtained as

$$Y_{U(k)}^{BP_1} = E(Y|\underset{\sim}{x})$$
$$= \int_0^\infty \left[\int_0^\infty \int_0^\infty \frac{\alpha \lambda^k}{\Gamma(k)} y^{\alpha k} e^{-\lambda y^\alpha} \pi(\alpha, \lambda|data) \, d\alpha \, d\lambda \right] dy.$$

Based on MC samples $\{(\alpha_i, \lambda_i); \ i = 1, 2, ..., M\}$ and by using the PDF of the gamma distribution, the simulation consistent based predictor $\hat{Y}_{U(k)}^{BP1}$ of $Y = Y_{U(k)}$ will be

$$\hat{Y}_{U(k)}^{BP_1} = \frac{1}{M} \sum_{i=1}^{M} \frac{\Gamma(k + \frac{1}{\alpha_i})}{\lambda_i^{\frac{1}{\alpha_i}} \Gamma(k)}. \tag{3.3}$$

Under the absolute error loss function L_2, the corresponding BP of $Y_{U(k)}$ denoted by $Y_{U(k)}^{BP_2}$, is the median of the posterior predictive density of $Y_{U(k)}$, Eq.(3.2), which is obtained by solving the following equation with respect to $Y_{U(k)}^{BP_2}$

$$\int_{Y_{U(k)}^{BP_2}}^{\infty} \left[\int_0^{\infty} \int_0^{\infty} \frac{\alpha \lambda^k}{\Gamma(k)} y^{\alpha k - 1} e^{-\lambda y^{\alpha}} \pi(\alpha, \lambda | data) \, d\alpha \, d\lambda \right] dy = \frac{1}{2}.$$

Based on MC samples $\{(\alpha_i, \lambda_i); \ i = 1, 2, ..., M\}$, the simulation consistent based predictor $\hat{Y}_{U(k)}^{BP_2}$ of $Y_{U(k)}$ can be obtained by solving the following equation with respect to $\hat{Y}_{U(k)}^{BP_2}$

$$\frac{1}{M} \sum_{i=1}^{M} \left[\frac{\gamma\left(k, \lambda_i (\hat{Y}_{U(k)}^{BP_2})^{\alpha_i}\right)}{\Gamma(k)} \right] = \frac{1}{2}. \tag{3.4}$$

Under the LINEX loss function L_3, the corresponding BP $Y_{U(k)}^{BP_3}$ of $Y = Y_{U(k)}$ for $1 \le k \le n$ can be obtained as

$$Y_{U(k)}^{BP3} = \left[E(Y^{-a^*} | data) \right]^{-\frac{1}{a^*}}$$

$$= \left[\int_0^{\infty} \left(\int_0^{\infty} \int_0^{\infty} \frac{\alpha \lambda^k}{\Gamma(k)} y^{\alpha k - a^* - 1} e^{-\lambda y^{\alpha}} \pi(\alpha, \lambda | data) \, d\alpha \, d\lambda \right) dy \right]^{-\frac{1}{a^*}}.$$

Based on MC samples $\{(\alpha_i, \lambda_i); \ i = 1, 2, ..., M\}$, the simulation consistent based predictor $\hat{Y}_{U(k)}^{BP_3}$ of $Y_{U(k)}^{BP_3}$ will be

$$\hat{Y}_{U(k)}^{BP_3} = \left[\frac{1}{M} \sum_{i=1}^{M} \frac{\lambda_i^{\frac{a^*}{\alpha_i}} \Gamma\left(k - \frac{a^*}{\alpha_i}\right)}{\Gamma(k)} \right]^{-\frac{1}{a^*}}. \tag{3.5}$$

To obtain prediction bounds on $Y = Y_{U(k)}$, for $1 \le k \le n$, we need the predictive distribution function of $Y = Y_{U(k)}$, which depends on the distribution function of $Y = Y_{U(k)}$. Based on Eq.(3.1), the PDF of $Y_{U(k)}$, the distribution function of $Y_{U(k)}$ can be written using the incomplete gamma function as follows:

$$G_{(k)}(y | \alpha, \lambda) = 1 - \frac{\gamma(k, \lambda y^{\alpha})}{\Gamma(k)}. \tag{3.6}$$

Arguments similar to those in subsection 3.1, the simulation consistent based predictor of the predictive distribution function of $Y_{U(k)}$, $G_{(k)}(y | \alpha, \lambda)$ based on the squared error loss function L_1 is

found to be

$$\hat{G}_{(k)}(y) = \frac{1}{M} \sum_{i=1}^{M} \left[1 - \frac{\gamma(k,\ \lambda_i y^{\alpha_i})}{\Gamma(k)} \right].$$

Under the absolute error loss function L_2, the following algorithm can be implemented to find the simulation consistent based predictor of $G_{(k)}(y|\alpha,\lambda)$.

Algorithm 2:

- **Step 1** Evaluate $G = G_{(k)}(y|\alpha,\lambda)$, Eq.(3.6), at each sample (α_i,λ_i) for $i = 1,2,...,M$, to get G_1, G_2, ... , G_M.
- **Step 2** Order G_1, G_2, ... , G_M as $G_{(1)} < G_{(2)} < ... < G_{(M)}$.
- **Step 3** The simulation consistent based predictor for $G_{(k)}(y|\alpha,\lambda)$ is given by

$$\hat{G}_{(k)}(y) = Median\left[G_{(1)},\ G_{(2)},\ ...\ ,\ G_{(M)} \right]$$

Under the LINEX loss function L_3 and Based on MC samples $\{(\alpha_i,\lambda_i);\ i = 1,2,...,M\}$, the simulation consistent based predictor of $G_{(k)}(y|\alpha,\lambda)$ is readily obtained as

$$\hat{G}_{(k)}(y) = \left[\frac{1}{M} \sum_{i=1}^{M} \left(1 - \frac{\gamma(k,\ \lambda_i y^{\alpha_i})}{\Gamma(k)} \right)^{-a*} \right]^{-\frac{1}{a*}}.$$

Under all different loss functions L_1, L_2 and L_3, the $(1-\beta)\%$ PI for $Y = Y_{U(k)}$, $1 \le k \le n$, can be obtained by solving the non-linear equations in Eq.(3.7) for the lower bound L and upper bound U:

$$\hat{G}_{(k)}(L) = \frac{\beta}{2}, \ \text{and} \ \hat{G}_{(k)}(U) = 1 - \frac{\beta}{2}. \tag{3.7}$$

As before, we need a suitable numerical method to solve these non-linear equations.

4. Numerical experiments and data analysis

In this section we conduct a simulation study to compute the different BPs of future records using one-sequence and two-sequence prediction problems. In both cases, we have assumed $\alpha = 2$, $\lambda = 1$ to generate record data from WE(α,λ). The first m observed records were generated by using the transformation:

$$X_{U(k)} = \left(\frac{\sum_{i=1}^{k} e(i)}{\lambda} \right)^{\frac{1}{\alpha}}, k = 1,2,...,m,$$

where $\{e(i), i \ge 0\}$ is a sequence of i.i.d Exp(1) [see [6], p.20]. We compute the BPs for the Weibull records, with respect to different loss functions: squared error (Sq. err.) (L_1), absolute error (Abs. err.) (L_2) and LINEX (L_3) with different choices of a^* : $0.1, 0.5, 1$. The prior $\pi_2(\alpha)$ on α can take any log-concave density on the support $(0,\infty)$. Here, we consider that the prior of α, has a gamma density function with the shape and scale parameters c and d, respectively. For the computations of BPs, we consider two types of prior for both α and λ: first prior is the non-informative prior, *i.e* $a = b = c = d = 0$, we call this prior as Prior 0, second prior is the informative prior, namely

$a = b = 1, c = 2, d = 1$, we call this prior as Prior 1. We have computed the PIs for the future Weibull records.

In Table 2, we present the simulated point predictors and PIs for the future *nth* record $X_{U(n)}, 1 \le m < n$ for the following cases of sample sizes $m = 6, m = 9$, and $m = 12$ and for all different loss functions L_1, L_2 and L_3 with different choices of $a^* : 0.1, 0.5, 1.0$. Based on MC samples $\{(\alpha_i, \lambda_i), i = 1, 2, ..., M\}$ with $M = 1000$ computed using the algorithm in Appendix C, the simulated predictors for the future *nth* record $X_{U(n)}, 1 \le m < n$, were computed by using the equations (2.3), (2.5) and (2.6), respectively. The 95% lower bound L and upper bound U of the PI for the future *nth* record were computed by solving the equations (2.10) and (2.11) with respect to L and U, respectively. In this table, the first three future *nth* records after the last observed record are only predicted. The mean square errors (MSEs) of the resulting predictors using Prior 0 and Prior 1 are reported in Table 3 based on $\ell = 1000$ runs are computed. The MSE of any BP (say $\underset{\sim}{\delta}(x)$) of the nth record, $X_{U(n)}$ is defined by

$$MSE(\underset{\sim}{\delta}(x)) = \frac{\sum_{i=1}^{\ell} \left(\underset{\sim}{\delta_i}(x) - X_{U(n)} \right)^2}{\ell},$$

where $\underset{\sim}{\delta_i}(x)$ is the predicted value of $X_{U(n)}$ based on the ith replicate.

In Table 4, we present the simulated predictors and PIs for the unobserved *kth* record $Y_{U(k)}, k = 1, ..., n$, based on observed record sample of size m, and for all different loss functions. Based on MC samples $\{(\alpha_i, \lambda_i), i = 1, 2, ..., M\}$ and $M = 1000$, the simulated predicted values for the unobserved *kth* record $Y_{U(k)}$, were computed under different loss functions, by using the equations (3.3), (3.4) and (3.5), respectively. The 95% lower bound L and upper bound U of the PI for the unobserved *kth* record $Y_{U(k)}$ were computed by solving the equations (3.7) with respect to L and U, respectively. In this table the smallest, middle or around, and the largest unobserved *kth* record $Y_{U(k)}$ are only predicted. The MSEs of these simulated predictors are also computed and presented in Table 5.

From Tables 2 and 4, it can be checked that the 95% PI for future record gets wide whenever the record to be predicted moves away. This is an expected observation since the variation of future record becomes high as it moves away. It is also observed from Table 2, the BP based on LINEX loss function provides shortest PIs comparing with the ones based on squared and absolute loss functions. Table 4 shows that the 95% Bayes PIs based on L_3 are almost similar. That is, the Bayes PIs based on two-sequence prediction case are robust with respect to the parameter a^* involved in LINEX loss function. For both prediction problems, it is also evident from Tables 3 and 5 the BPs obtained by using Prior 1 (informative prior) perform well comparing with the corresponding ones obtained by using Prior 0 (non-informative prior).

Example (real data): In this example we analyze the total seasonal annual rainfall (in inches) recorded at Loss Angeles Civic Center during 132 years, from 1878 to 2009 (season July 1 - June 30). The data set can be obtained from the loss Angeles Civic Website: http://www.laalmanac.com/weather/we13.htm. For the complete data set, we have computed the MLEs of α and λ and they are 2.2438 and 0.0018, respectively. The corresponding Kolmogorov-Smirnov (KS) distance becomes 0.0939 and the associated p-value is 0.1949. Therefore the KS indicates that Weibull distribution can be used to analyze this

rainfall data. We used the upper records from 1930 to 2009 which were as follows:

$$12.54, 16.93, 21.66, 22.41, 23.43, 32.76, 33.44, 37.96.$$

From the observed data, it is also possible to estimate the shape of the hazard function. A tool called scaled TTT transform and its empirical version are considered in this context. If a family has a survival function $S(y) = 1 - F(y)$, then the scaled TTT transform is $g(u) = H^{-1}(u)/H^{-1}(1)$ for $0 < u < 1$, where $H^{-1}(u) = \int_0^{F^{-1}(u)} S(y)\,dy$. The empirical version of the scaled TTT transform is given by

$$g(r/n) = H_n^{-1}(r/n)/H_n^{-1}(1) = \frac{\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}}{\sum_{i=1}^{n} X_{i:n}},$$

where $X_{i:n}$ is the ith order statistic from a sample of size n. It has been shown by [1] the scaled TTT transform is convex (concave) if the hazard rate is decreasing (increasing) and for bathtub (unimodal) shaped hazard rate, the scaled TTT transform is first convex (concave) and then concave (convex). In this example, the scaled TTT transform of the rainfall data presented in Fig. 1, shows the scaled TTT transform is almost concave; we therefore conclude that the hazard function is increasing function. This indicates that the WE model may provide a reasonable fit to the above data.

Now, we compute the BEs with respect to different loss functions: squared error (Sq. err.), absolute error (Abs. err.) and LINEX function with different choices of a^*: $0.1, 0.5, 1.0$ when Prior 1 is used. The BEs of λ and α are computed and presented as follows:

Table 1. The BEs of λ and α when Prior 1 is used.

	Sq. err.	Abs. err.	$a^* = 0.1$	$a^* = 0.5$	$a^* = 1.0$
	BE_1	BE_2	BE_3	BE_4	BE_5
α	1.8160	1.8249	1.8155	1.8151	1.8134
λ	0.0110	0.0103	0.0101	0.0094	0.0060

All the estimates are quite close to each other. We obtain the 95% credible intervals for α and λ and they are $(1.7187, 1.8592)$ and $(0.0045, 0.0206)$, respectively. Also, we consider the prediction of the 9th, 10th and 11th future records. The simulated predictors and the 95% PIs of the 9th, 10th and 11th future records are presented in Table 6. It is observed that all predicted values, with respect to different loss functions, are all ordered and fall in their corresponding predictive intervals.

To study the sensitivity of our results to variations in the specification of prior parameters, further MC simulations were performed using two additional proper priors and two improper priors. For the first proper prior (π_1), we assume that the prior mean of α is equal to 2 and its standard deviation is equal to 1, and that for λ, the prior mean is 1 with a standard deviation of 0.33. This corresponds to $a = 4, b = 2, c = 9$, and $d = 9$. For the second proper prior (π_2), we assume that only small amount of prior information is available and assign small integer values to a, b, c, and d. For example, let us assume that $a = 3, b = 1$ and $c = 3, d = 1$. Priors π_3 and π_4 are limiting improper priors obtained by letting $a = 0.5, b = 0$

and $c = 0.5, d = 0$ for the first, $a = 1, b = 0$ and $c = 0, d = 0$ for the second. Table 7 shows the MC based BEs for α and λ, and the PIs for the next record statistic $X_{U(9)}$. It is evident from Table 7 that our results are not very sensitive to the assumed values of the prior parameters. Note that PIs corresponding to all four priors include the predicted value of the next record statistic, $X_{U(9)}$.

5. Conclusions

Motivated by the importance of the prediction problem of future observations, various inferential methods have been discussed in the literature. We have proposed a Bayesian approach to predict the future records based on some observed ones in one-sequence and two-sequence prediction problems. In Bayesian analysis, the so obtained predictors depends on the prior distribution and also on the error loss function chosen. For this, our proposed priors are quite flexible in nature and Jeffrey's prior can be obtained as a special case of the proposed priors. Additionally, we have used different error loss functions. We opt for sampling simulation procedures, namely, the Monte Carlo samplers to generate samples from the predictive distributions and then obtain different BPs as well as PIs of the future records. In this context, these prediction problems are not considered previously in the literature. Our methods of prediction were simple and quite useful since they allow us to assume priors with general set-up and different error loss functions. The simulation algorithm by [11] and [12] can be used effectively for generating random variates with a log-concave density. In the sense of prediction interval length, it is observed that the BPs under the LINEX loss function perform well when compared to the BPs under other loss functions. Moreover, it is evident that, the BPs are not very sensitive to the assumed values of the prior parameters (Prior 0 and Prior 1).
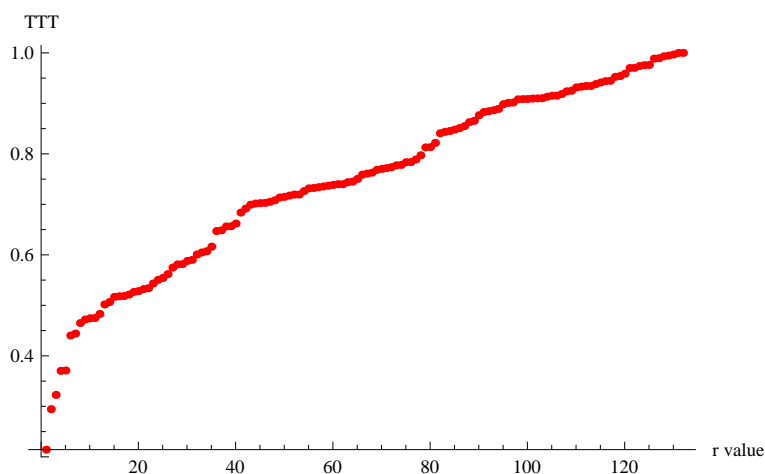

Fig. 1. The empirical scaled TTT transform of the rainfall data.

Table 2. Simulated predictors and PIs for future records $X_{U(n)}$ based on m observed records ($n \geq m+1$).

Number of observed records	$X_{U(n)}$	Loss function	Prior 0		Prior 1	
			Predicted values	95% PIs	Predicted values	95% PIs
$m = 6$		L_1	2.4335	(2.1822, 3.2599)	2.3578	(2.1814, 2.8334)
		L_2	2.3341	(2.1838, 3.0877)	2.3040	(2.1815, 2.8361)
	$X_{U(7)}$	$L_3(a^* = 0.1)$	2.4149	(2.1822, 2.8702)	2.3510	(2.1814, 2.7635)
		$L_3(a^* = 0.5)$	2.4091	(2.1822, 2.7883)	2.3487	(2.1814, 2.7418)
		$L_3(a^* = 1.0)$	2.4025	(2.1822, 2.7121)	2.3459	(2.1814, 2.7001)
		L_1	2.6777	(2.2255, 3.9064)	2.5253	(2.2202, 3.1313)
		L_2	2.5447	(2.2444, 3.4827)	2.4720	(2.2212, 3.1296)
	$X_{U(8)}$	$L_3(a^* = 0.1)$	2.6393	(2.2252, 3.1562)	2.5134	(2.2201, 3.0087)
		$L_3(a^* = 0.5)$	2.6278	(2.2251, 3.0163)	2.5094	(2.2201, 2.9716)
		$L_3(a^* = 1.0)$	2.6146	(2.2250, 2.8975)	2.5044	(2.2201, 2.8300)
		L_1	2.9121	(2.2940, 4.5039)	2.6816	(2.2842, 3.3758)
		L_2	2.7454	(2.3465, 3.8041)	2.6294	(2.2866, 3.3679)
	$X_{U(9)}$	$L_3(a^* = 0.1)$	2.8525	(2.2931, 3.3831)	2.6659	(2.2840, 3.2047)
		$L_3(a^* = 0.5)$	2.8351	(2.2928, 3.1909)	2.6604	(2.2839, 3.1534)
		$L_3(a^* = 1.0)$	2.8154	(2.2925, 3.1598)	2.6538	(2.2838, 3.0967)
$m = 9$		L_1	3.2205	(2.8408, 4.5014)	3.0904	(2.8388, 3.7810)
		L_2	3.0775	(2.8401, 3.9837)	3.0118	(2.8389, 3.7386)
	$X_{U(10)}$	$L_3(a^* = 0.1)$	3.1913	(2.8408, 4.0579)	3.0797	(2.8388, 3.7006)
		$L_3(a^* = 0.5)$	3.1818	(2.8408, 3.9425)	3.0760	(2.8388, 3.6734)
		$L_3(a^* = 1.0)$	3.1705	(2.8408, 3.8166)	3.0717	(2.8388, 3.6407)
		L_1	3.6140	(2.9102, 5.4121)	3.3363	(2.8937, 4.2349)
		L_2	3.4272	(2.9072, 4.5780)	3.2557	(2.8958, 4.1549)
	$X_{U(11)}$	$L_3(a^* = 0.1)$	3.5534	(2.9100, 4.6580)	3.3171	(2.8936, 4.0894)
		$L_3(a^* = 0.5)$	3.5333	(2.9099, 4.4367)	3.3105	(2.8936, 4.0403)
		$L_3(a^* = 1.0)$	3.5083	(2.9097, 4.2067)	3.3025	(2.8936, 3.9817)
		L_1	4.0124	(3.0248, 6.4919)	3.5718	(2.9855, 4.6180)
		L_2	3.7837	(3.0238, 5.1028)	3.4909	(2.9923, 4.5011)
	$X_{U(12)}$	$L_3(a^* = 0.1)$	3.9183	(3.0238, 5.1759)	3.5456	(2.9852, 4.4097)
		$L_3(a^* = 0.5)$	3.8862	(3.0234, 4.8417)	3.5366	(2.9851, 4.3393)
		$L_3(a^* = 1.0)$	3.8435	(3.0230, 4.5162)	3.5256	(2.9850, 4.2555)
$m = 12$		L_1	3.6338	(3.4261, 4.2522)	3.5929	(3.4254, 4.0568)
		L_2	3.5619	(3.4260, 4.1535)	3.5399	(3.4251, 3.9942)
	$X_{U(13)}$	$L_3(a^* = 0.1)$	3.6266	(3.4261, 4.1116)	3.5886	(3.4254, 4.0059)
		$L_3(a^* = 0.5)$	3.6241	(3.4261, 4.0684)	3.5871	(3.4254, 3.9886)
		$L_3(a^* = 1.0)$	3.6212	(3.4261, 4.0200)	3.5853	(3.4254, 3.9676)
		L_1	3.8415	(3.4674, 4.6978)	3.7583	(3.4617, 4.3663)
		L_2	3.7577	(3.4679, 4.5185)	3.7032	(3.3885, 4.2699)
	$X_{U(14)}$	$L_3(a^* = 0.1)$	3.8272	(3.4673, 4.4336)	3.7503	(3.4617, 4.2734)
		$L_3(a^* = 0.5)$	3.8224	(3.4673, 4.3541)	3.7475	(3.4617, 4.2418)
		$L_3(a^* = 1.0)$	3.8165	(3.4672, 4.2685)	3.7441	(3.4616, 4.2035)
		L_1	4.0445	(3.5360, 5.0916)	3.9178	(3.5228, 4.6289)
		L_2	3.9513	(3.5391, 4.8325)	3.8621	(3.5159, 4.5028)
	$X_{U(15)}$	$L_3(a^* = 0.1)$	4.0234	(3.5356, 4.7037)	3.9066	(3.5227, 4.4953)
		$L_3(a^* = 0.5)$	4.0163	(3.5355, 4.5882)	3.9026	(3.5226, 4.4497)
		$L_3(a^* = 1.0)$	4.0076	(3.5353, 4.4688)	3.8978	(3.5225, 4.3945)

Table 3. MSEs of the predictors of future records based on one-sequence prediction problem.

m	n	Loss Function	Prior 0	Prior 1
6	7	L_1	0.3551	0.2569
		L_2	0.3440	0.2305
		$L_3(a^* = 0.1)$	0.3644	0.2507
		$L_3(a^* = 0.5)$	0.3622	0.2488
		$L_3(a^* = 1)$	0.3601	0.2467
	8	L_1	0.4661	0.3452
		L_2	0.4531	0.3032
		$L_3(a^* = 0.1)$	0.4634	0.3284
		$L_3(a^* = 0.5)$	0.4703	0.3242
		$L_3(a^* = 1)$	0.4627	0.3196
	9	L_1	0.4625	0.4011
		L_2	0.4223	0.3730
		$L_3(a^* = 0.1)$	0.5556	0.4017
		$L_3(a^* = 0.5)$	0.4952	0.3940
		$L_3(a^* = 1)$	0.4501	0.3853
9	10	L_1	0.2396	0.2281
		L_2	0.2278	0.2196
		$L_3(a^* = 0.1)$	0.2377	0.2272
		$L_3(a^* = 0.5)$	0.2371	0.2269
		$L_3(a^* = 1)$	0.2364	0.2265
	11	L_1	0.3027	0.2155
		L_2	0.2901	0.2000
		$L_3(a^* = 0.1)$	0.3002	0.2118
		$L_3(a^* = 0.5)$	0.2994	0.2105
		$L_3(a^* = 1)$	0.2985	0.2091
	12	L_1	0.4022	0.2580
		L_2	0.3718	0.2449
		$L_3(a^* = 0.1)$	0.3908	0.2548
		$L_3(a^* = 0.5)$	0.3873	0.2537
		$L_3(a^* = 1)$	0.3834	0.2525
12	13	L_1	0.3870	0.2577
		L_2	0.3728	0.2498
		$L_3(a^* = 0.1)$	0.3854	0.2571
		$L_3(a^* = 0.5)$	0.3848	0.2569
		$L_3(a^* = 1)$	0.3842	0.2566
	14	L_1	0.2821	0.2458
		L_2	0.2701	0.2387
		$L_3(a^* = 0.1)$	0.2796	0.2448
		$L_3(a^* = 0.5)$	0.2788	0.2445
		$L_3(a^* = 1)$	0.2779	0.2441
	15	L_1	0.3173	0.2326
		L_2	0.3078	0.2220
		$L_3(a^* = 0.1)$	0.3151	0.2300
		$L_3(a^* = 0.5)$	0.3143	0.2291
		$L_3(a^* = 1)$	0.3135	0.2281

Table 4. Simulated predictors and PIs for unobserved records $Y_{U(k)}, k = 1, ..., n$, based on another independent observed record sample of size m.

m	n	$Y_{U(k)}$	Loss function	Prior 0 Predicted values	Prior 0 95% PIs	Prior 1 Predicted values	Prior 1 95% PIs
6	4		L_1	0.8372	(0.0696, 1.9504)	0.8880	(0.1880, 1.8580)
			L_2	0.7895	(0.1082, 1.4858)	0.8417	(0.1736, 1.3285)
		$Y_{U(1)}$	$L_3(a^* = 0.1)$	0.7049	(0.1171, 1.9701)	0.7551	(0.1908, 1.8611)
			$L_3(a^* = 0.5)$	0.6359	(0.1273, 1.9776)	0.6974	(0.1919, 1.8623)
			$L_3(a^* = 1.0)$	0.5243	(0.1381, 1.9874)	0.6127	(0.1933, 1.8638)
			L_1	1.4669	(0.4917, 3.0124)	1.5775	(0.7786, 2.5618)
			L_2	1.3936	(0.5747, 2.1219)	1.5452	(0.7285, 2.0040)
		$Y_{U(3)}$	$L_3(a^* = 0.1)$	1.3531	(0.6719, 3.0864)	1.5027	(0.8075, 2.5700)
			$L_3(a^* = 0.5)$	1.2109	(0.7287, 3.1178)	1.4745	(0.8190, 2.5731)
			$L_3(a^* = 1.0)$	1.2565	(0.8071, 3.1608)	1.4381	(0.8342, 2.5770)
			L_1	1.7330	(0.6956, 3.4447)	1.8141	(0.9984, 2.8124)
			L_2	1.6792	(0.7720, 2.3517)	1.7823	(0.9392, 2.2048)
		$Y_{U(4)}$	$L_3(a^* = 0.1)$	1.6252	(0.9175, 3.5629)	1.7480	(1.0440, 2.8232)
			$L_3(a^* = 0.5)$	1.5866	(1.0042, 3.6164)	1.7233	(1.0625, 2.8273)
			$L_3(a^* = 1.0)$	1.5381	(1.1335, 3.6930)	1.6920	(1.2876, 2.8326)
9	6		L_1	0.8228	(0.0970, 2.1196)	0.8910	(0.1394, 2.0942)
			L_2	0.7234	(0.1018, 1.8077)	0.8242	(0.1364, 1.7410)
		$Y_{U(1)}$	$L_3(a^* = 0.1)$	0.6213	(0.1074, 2.1320)	0.7212	(0.1414, 2.0982)
			$L_3(a^* = 0.5)$	0.5328	(0.1105, 2.1367)	0.6430	(0.1422, 2.0997)
			$L_3(a^* = 1.0)$	0.3965	(0.1142, 2.1428)	0.5254	(0.1432, 2.1016)
			L_1	1.6811	(0.6191, 3.2753)	1.7497	(0.7560, 3.0590)
			L_2	1.5893	(0.6324, 2.6786)	1.6948	(0.7487, 2.5779)
		$Y_{U(3)}$	$L_3(a^* = 0.1)$	1.5312	(0.6975, 3.3164)	1.6374	(0.7832, 3.0708)
			$L_3(a^* = 0.5)$	1.4751	(0.7290, 3.3336)	1.5949	(0.7944, 3.0754)
			$L_3(a^* = 1.0)$	1.4030	(0.7744, 3.3571)	1.5401	(0.8099, 3.0815)
			L_1	2.5706	(1.2772, 4.4939)	2.5864	(1.4669, 4.0211)
			L_2	2.4617	(1.3010, 3.5500)	2.5315	(1.4721, 3.3875)
		$Y_{U(6)}$	$L_3(a^* = 0.1)$	2.4324	(1.4788, 4.5927)	2.4962	(1.5514, 4.0486)
			$L_3(a^* = 0.5)$	2.3829	(1.5882, 4.6397)	2.4631	(1.5906, 4.0602)
			$L_3(a^* = 1.0)$	2.3213	(1.7899, 4.7092)	2.4215	(1.8531, 4.0761)
12	8		L_1	0.7924	(0.1076, 1.9409)	0.8857	(0.1501, 1.8947)
			L_2	0.7114	(0.1081, 1.8183)	0.8207	(0.1549, 1.7616)
		$Y_{U(1)}$	$L_3(a^* = 0.1)$	0.6172	(0.1136, 1.9470)	0.7229	(0.1521, 1.7976)
			$L_3(a^* = 0.5)$	0.5407	(0.1156, 1.9492)	0.6516	(0.1528, 1.7987)
			$L_3(a^* = 1.0)$	0.4240	(0.1179, 1.9521)	0.5451	(0.1538, 1.7603)
			L_1	1.8729	(0.8578, 3.2779)	1.9659	(1.0053, 3.1741)
			L_2	1.8051	(0.8352, 3.0217)	1.9226	(1.0528, 3.0003)
		$Y_{U(4)}$	$L_3(a^* = 0.1)$	1.7609	(0.9372, 3.3000)	1.8783	(1.0446, 3.1839)
			$L_3(a^* = 0.5)$	1.7194	(0.9695, 3.3083)	1.8455	(1.0606, 3.1876)
			$L_3(a^* = 1.0)$	1.6668	(1.0121, 3.3189)	1.8038	(1.0821, 3.1924)
			L_1	2.8015	(1.6010, 4.4560)	2.8449	(1.7839, 4.1588)
			L_2	2.7223	(1.5448, 4.8359)	2.8004	(1.8942, 3.9318)
		$Y_{U(8)}$	$L_3(a^* = 0.1)$	2.7000	(1.8028, 4.5046)	2.7740	(1.8933, 4.1797)
			$L_3(a^* = 0.5)$	2.6635	(1.8998, 4.5230)	2.7481	(1.9419, 4.1880)
			$L_3(a^* = 1.0)$	2.6179	(2.0305, 4.5465)	2.7156	(2.1122, 4.1992)

Table 5. MSEs of the predictors of unobserved records based on two-sequence prediction problem.

m	n	$Y_{U(k)}$	Loss Function	Prior 0	Prior 1
6	3	$Y_{U(1)}$	L_1	0.1606	0.0518
			L_2	0.1907	0.0446
			$L_3(a^* = 0.1)$	0.1898	0.0386
			$L_3(a^* = 0.5)$	0.2195	0.0353
			$L_3(a^* = 1)$	0.2596	0.0339
		$Y_{U(2)}$	L_1	0.1830	0.0626
			L_2	0.3151	0.0558
			$L_3(a^* = 0.1)$	0.3041	0.0535
			$L_3(a^* = 0.5)$	0.3274	0.0523
			$L_3(a^* = 1)$	0.3771	0.0531
		$Y_{U(3)}$	L_1	0.2927	0.0790
			L_2	0.4185	0.0603
			$L_3(a^* = 0.1)$	0.4755	0.0646
			$L_3(a^* = 0.5)$	0.5972	0.0604
			$L_3(a^* = 1)$	0.6474	0.0707
9	6	$Y_{U(1)}$	L_1	0.1582	0.0306
			L_2	0.1189	0.0385
			$L_3(a^* = 0.1)$	0.1794	0.0371
			$L_3(a^* = 0.5)$	0.1868	0.0406
			$L_3(a^* = 1)$	0.2144	0.0490
		$Y_{U(3)}$	L_1	0.1703	0.0742
			L_2	0.1975	0.0775
			$L_3(a^* = 0.1)$	0.1988	0.0759
			$L_3(a^* = 0.5)$	0.2108	0.0773
			$L_3(a^* = 1)$	0.2281	0.0797
		$Y_{U(6)}$	L_1	0.2266	0.1458
			L_2	0.2458	0.1300
			$L_3(a^* = 0.1)$	0.2506	0.1280
			$L_3(a^* = 0.5)$	0.2660	0.1227
			$L_3(a^* = 1)$	0.3138	0.1170
12	9	$Y_{U(1)}$	L_1	0.1145	0.0314
			L_2	0.1324	0.0365
			$L_3(a^* = 0.1)$	0.1246	0.0344
			$L_3(a^* = 0.5)$	0.1295	0.0363
			$L_3(a^* = 1)$	0.1392	0.0408
		$Y_{U(5)}$	L_1	0.2054	0.1157
			L_2	0.2247	0.1123
			$L_3(a^* = 0.1)$	0.2261	0.1112
			$L_3(a^* = 0.5)$	0.2343	0.1101
			$L_3(a^* = 1)$	0.2453	0.1092
		$Y_{U(9)}$	L_1	0.3093	0.1147
			L_2	0.3050	0.1148
			$L_3(a^* = 0.1)$	0.3077	0.1130
			$L_3(a^* = 0.5)$	0.3085	0.1130
			$L_3(a^* = 1)$	0.3107	0.1133

Table 6. Simulated predictors and PIs for the 9th, 10th and 11th future records.

Number of observed records	$X_{U(n)}$	Loss function	Predicted values	95% PIs
$m = 8$	$X_{U(9)}$	L_1	42.76	$(38.07, 57.32)$
		L_2	41.05	$(38.06, 57.05)$
		$L_3(a^* = 0.1)$	42.45	$(38.07, 53.48)$
		$L_3(a^* = 0.5)$	42.35	$(38.07, 52.31)$
		$L_3(a^* = 1.0)$	42.23	$(38.07, 51.04)$
	$X_{U(10)}$	L_1	47.56	$(38.96, 68.48)$
		L_2	45.44	$(38.88, 67.92)$
		$L_3(a^* = 0.1)$	46.94	$(38.96, 61.07)$
		$L_3(a^* = 0.5)$	46.74	$(38.96, 58.87)$
		$L_3(a^* = 1.0)$	46.51	$(38.96, 56.61)$
	$X_{U(11)}$	L_1	52.36	$(40.44, 78.73)$
		L_2	49.88	$(40.22, 77.63)$
		$L_3(a^* = 0.1)$	51.44	$(40.43, 67.60)$
		$L_3(a^* = 0.5)$	51.14	$(40.43, 64.35)$
		$L_3(a^* = 1.0)$	50.78	$(40.42, 61.17)$

Table 7. Bayesian analysis using different priors.

Prior		L_1	L_2	$L_3(a^* = 0.1)$	$L_3(a^* = 0.5)$	$L_3(a^* = 1.0)$
π_1	$\hat{\alpha}$	1.8193	1.8241	1.8187	1.8183	1.8160
	$\hat{\lambda}$	0.0115	0.0122	0.0104	0.0114	0.0120
	Predicted Value	42.25	41.79	42.01	41.93	41.84
	95% PI of $X_{U(9)}$	$(38.06, 50.77)$	$(38.08, 50.40)$	$(38.06, 48.11)$	$(38.06, 47.24)$	$(38.06, 47.22)$
π_2	$\hat{\alpha}$	1.7965	1.8029	1.7960	1.7956	1.7934
	$\hat{\lambda}$	0.0121	0.0123	0.0120	0.0127	0.0129
	Predicted Value	42.06	41.01	41.92	41.88	41.83
	95% PI of $X_{U(9)}$	$(38.03, 52.05)$	$(38.03, 52.52)$	$(38.03, 50.25)$	$(38.03, 48.15)$	$(38.03, 48.05)$
π_3	$\hat{\alpha}$	1.8304	1.8312	1.8304	1.8303	1.8322
	$\hat{\lambda}$	0.0109	0.0113	0.0105	0.0116	0.0119
	Predicted Value	42.03	41.98	41.90	41.85	41.80
	95% PI of $X_{U(9)}$	$(38.03, 54.05)$	$(38.03, 54.53)$	$(38.03, 53.25)$	$(38.03, 52.15)$	$(38.03, 50.05)$
π_4	$\hat{\alpha}$	1.8812	1.8791	1.8706	1.8690	1.8729
	$\hat{\lambda}$	0.0114	0.0108	0.0109	0.0114	0.0116
	Predicted Value	42.56	41.31	42.38	42.32	42.25
	95% PI of $X_{U(9)}$	$(38.04, 55.26)$	$(38.02, 55.06)$	$(38.04, 54.40)$	$(38.04, 53.49)$	$(38.04, 52.43)$

References

[1] M. V. Aarset, How to identify a bathtub hazard rate, *IEEE Transactions on Reliability* **36** (1987), 106 – 108.

[2] J. Ahmadi and M. Doostparast, Bayesian estimation and prediction for some life distributions based on record values, *Statistical Papers* **47** (2006), 373–392.

[3] M. Ahsanullah, Linear prediction of record values for the two parameter exponential distriburion, *Annals of the Institute of Statistical Mathematics* **32** (1980), 363–368.

[4] M. Ahsanullah, *Record values – theory and applications*, (University Press of America Inc., New York, 2004).

[5] E.K. Al-Hussaini, Predicting observable from a general class of distributions, *Journal of Statistical Planning and Inference* **79** (1999), 79–81.

[6] B.C. Arnold, N. Balakrishnan and H.N. Nagaraja, *Records*, (Jon Weily, New York, 1998).

[7] A. Asgharzadeh and A. Fallah, Estimation and Prediction for Exponentiated Family of Distributions Based on Records, *Communications in Statistics - Theory and Methods* **40(1)** (2010), 68–83.

[8] A. Asgharzadeha, R. Valiollahib and D. Kundu, Prediction for future failures in Weibull distribution under hybrid censoring, *Journal of Statistical Computation and Simulation* **85(4)** (2015), 824–838.

[9] J.O. Berger and D. Sun, Bayesian analysis for the Poly-Weibull distribution, *Journal of the American Statistical Association* **88** (1993), 1412–1418.

[10] P. Dellaportas and D. Wright, Numerical prediction for the two-parameter Weibull distribution, *The Statistician* **40** (1991), 365–372.

[11] L. Devroye, A simple algorithm for generating random variates with a log-concave density, *Computing* **33** (1984), 247–257.

[12] L. Devroye, *Non-Uniform Random Variate Generation*, (Springer, Germany, 1986).

[13] H. K. Hsieh, Prediction intervals for Weibull order statistics, *Statistica Sinica* **7** (1997), 1039–1051

[14] S. R. Huanga and S. J. Wu, Bayesian estimation and prediction forWeibull model with progressive censoring, *Journal of Statistical Computation and Simulation* **82(11)** (2012), 1607–1620.

[15] Z. F. Jaheen, Empirical Bayes analysis of record statistics based on LINEX and quadratic loss functions, *Computers and Mathematics with Applications* **47** (2004), 947–954.

[16] N.L. Johnson, S. Kotz and N. Balakrishnan, *Continuous Univariate Distribution* 1, 2nd Ed., (Wiley, New York, 1995).

[17] V. B. Nevzorov, *Records: Mathematical Theory* (English Translation), (American Mathematical Society, Providence, Rhode Island, 2000).

[18] C. Ren, D. Sun and D. K. Dey, Bayes and frequentist estimation and prediction for exponential distribution, *Journal of Statistical Planning and Inference* **136** (2006), 2873–2897.

[19] R. Soland, Bayesian Analysis of the Weibull Process With Unknown Scale and Shape Parameters, *IEEE Transactions on Reliability Analysis* **18** (1969), 181–184.

[20] A. A. Soliman, A. H. Abd Ellah and K. S. Sultan, Comparisons of estimates using record statistics from Weibull model: Bayesian and non-Bayesian approaches, *Computational Statistics & Data Analysis* **51** (2006), 2065–2077.

[21] K. S. Sultan, G. R. Al-Dayian and H. H. Mohammad, Estimation and prediction from gamma distribution based on record values, *Computational Statistics & Data Analysis* **52(3)** (2008), 1430–1440.

[22] M. Teimouri and A. K. Gupta, On the Weibull record statistics and associated inferences, *Statistica* **72(2), LXXII(2)** (2012), 145–162.

[23] H. R. Varian, A Bayesian approach to real estate assessment, *Studies in Bayesian Econometrics and Statistics in Honor of L.J. Savage*, (Amsterdam, North Holland, 1975), 195–208.

Appendix A.   (BEs of α and λ)

Let $X_{U(1)}, X_{U(2)}, ..., X_{U(n)}$ be the first n upper record values arising from a sequence of iid WE(α, λ) with PDF and CDF given respectively in (1.1) and (1.2). The likelihood function

of this sample [see [6] and [4]] is

$$L(\alpha, \lambda | data) = \alpha^n \lambda^n e^{-\lambda x_{U(n)}^\alpha} \prod_{i=1}^{n} x_{U(i)}^{\alpha-1}. \tag{A.1}$$

First, consider the case when the shape parameter α is known. By (A.1) and the prior density Eq.(1.3), the posterior density of λ given α and data is $Gamma(a+n, b+x_{U(n)}^\alpha)$ which has the form

$$\pi_1(\lambda | \alpha, data) = \frac{(b+x_{U(n)}^\alpha)^{a+n}}{\Gamma(a+n)} \lambda^{a+n-1} e^{-\lambda(b+x_{U(n)}^\alpha)}. \tag{A.2}$$

Under the squared error loss function L_1, the BE $\hat{\lambda}_{B_1}$ of λ is

$$\hat{\lambda}_{B_1} = E_{posterior}(\lambda | \alpha, data) = \frac{a+n}{b+x_{U(n)}^\alpha}.$$

Under Jeffrey's prior ($a = b = 0$), the BE under the loss function L_1, $\widehat{\lambda}_{B_1}$ is same as the MLE of λ. For the absolute error loss function L_2, the BE of λ is median of the posterior density (say $\hat{\lambda}_{B_2}$). It can be obtained by solving the following equation in w, numerically:

$$\gamma\left(a+n, (b+x_{U(n)}^\alpha)w\right) - \frac{\Gamma(a+n)}{2} = 0.$$

Under the LINEX loss function L_3 and for any given $a^* < a+n$, the BE $\hat{\lambda}_{B_3}$ of λ can be obtained by using the PDF of the gamma distribution as follows:

$$\hat{\lambda}_{B_3} = \left[E_{posterior}[\lambda^{-a^*} | data]\right]^{-\frac{1}{a^*}} = \left[\frac{\Gamma(n+a-a^*)}{\Gamma(a+n)}\right]^{-\frac{1}{a^*}} \frac{1}{b+x_{U(n)}^\alpha}.$$

For unknown α and λ and based on the prior distributions $\pi_1(\lambda | a, b)$ and $\pi_2(\alpha)$, the posterior distribution of α and λ is defined by

$$\pi(\alpha, \lambda | data) = \frac{L(\alpha, \lambda | data) . \pi_1(\lambda | \alpha, a, b)\pi_2(\alpha)}{\int_0^\infty \int_0^\infty L(\alpha, \lambda | data) . \pi_1(\lambda | \alpha, a, b)\pi_2(\alpha) d\alpha d\lambda}. \tag{A.3}$$

Under L_1, L_2 and L_3, the BE of any function α and λ(say $\theta = g(\alpha, \lambda)$) can be written, respectively, as

$$\hat{\theta}_{B_1} = \int_0^\infty \int_0^\infty \theta \pi(\alpha, \lambda | data) d\alpha d\lambda, \quad \hat{\theta}_{B_2} = Med_{posterior}(\theta | data),$$

and

$$\hat{\theta}_{B3} = \left[\int_0^\infty \int_0^\infty \theta^{-a^*} \pi(\alpha, \lambda | data) d\alpha d\lambda\right]^{-\frac{1}{a^*}}.$$

Appendix B. (Log-concavity of $\pi_2(\alpha|data)$)

The BEs $\hat{\theta}_{B_1}$, $\hat{\theta}_{B_2}$ and $\hat{\theta}_{B_3}$, under different loss functions can't be obtained in closed forms. Here we propose to use sampling based technique to generate MC samples from the posterior distribution $\pi(\alpha, \lambda|data)$. For this, we state the following lemma:

Lemma B.1. *The conditional PDF of α given the data is given by*

$$\pi_2(\alpha|data) \propto \pi_2(\alpha) \alpha^n \prod_{i=1}^{n} x_{U(i)}^{\alpha-1} \times \frac{1}{\left(b + x_{U(n)}^{\alpha}\right)^{a+n}}, \tag{B.1}$$

and it is log-concave if $\pi_2(\alpha)$ is log-concave.

Proof. From the posterior distribution of α and λ, Eq. (A.3), we have

$$\pi(\alpha, \lambda|data) \propto \pi_2(\alpha) \alpha^n \prod_{i=1}^{n} x_{U(i)}^{\alpha-1} \lambda^{a+n-1} e^{-\lambda(b+x_{U(n)}^{\alpha})}.$$

Therefore, the PDF of α given data is

$$\pi_2(\alpha|data) \propto \pi_2(\alpha) \alpha^n \prod_{i=1}^{n} x_{U(i)}^{\alpha-1} \times \frac{1}{\left(b + x_{U(n)}^{\alpha}\right)^{a+n}}.$$

To show the log-concavity of $\pi_2(\alpha|data)$, let us consider

$$\ln \pi_2(\alpha|data) = C + \ln \pi_2(\alpha) + n \ln(\alpha) + (\alpha - 1) \sum_{i=1}^{n} \ln x_{U(i)} - (a+n) \ln(b + x_{U(n)}^{\alpha}),$$

where C is some constant. Now the second derivative can be written as

$$(\ln \pi_2(\alpha|data))'' = (\ln \pi_2(\alpha))'' - \frac{n}{\alpha^2} - (a+n) \frac{b \, x_{U(n)}^{\alpha} [\ln x_{U(n)}]^2}{[b + x_{U(n)}^{\alpha}]^2}.$$

Since $\pi_2(\alpha)$ is log-concave by the assumption, we have $(\ln \pi_2(\alpha))'' < 0$ and thus $(\ln \pi_2(\alpha|data))'' < 0$, for $a, b > 0$. It follows that $\pi_2(\alpha|data)$ is log-concave density. $\square$

Appendix C. (Generation of MC samples from $\pi(\alpha, \lambda)$)

The method proposed by [11] can be used to generate α from the distribution with PDF $\pi_2(\alpha|data)$. We then can generate λ based on the fact that λ given α and data is $Gamma(a+n, b+x_{U(n)}^{\alpha})$. Now, it is possible to generate MC samples from the posterior distribution of α and λ, Eq.(A.3), and then use these samples to obtain the BEs of any function of α and λ, $\theta = g(\alpha, \lambda)$.

Algorithm for obtaining sample-based estimates:

- **Step 1:** Generate α from the log-concave density function $\pi_2(\alpha|data)$, Eq.(B.1), using the method proposed by Devroye (1984).
- **Step 2:** For each α, generate λ from the posterior density function of λ given α and data, $\pi_1(\lambda|\alpha, data)$, Eq.(A.2).
- **Step 3:** Repeat step 1 and step 2, M times and obtain MC samples $\{(\alpha_i, \lambda_i); i = 1, ..., M\}$.
- **Step 4:** Obtain the BE of $\theta = g(\alpha, \lambda)$ with respect to the squared, absolute and LINEX error loss functions as

$$\widehat{\theta}_{B_1} = \frac{1}{M}\sum_{i=1}^{M} g(\alpha_i, \lambda_i), \quad \widehat{\theta}_{B_2} = Median\{\theta_{(1)}, \theta_{(2)}, ..., \theta_{(M)}\},$$

and

$$\widehat{\alpha}_{B_3} = \left[\frac{1}{M}\sum_{i=1}^{M} \frac{1}{g^{a^*}(\alpha_i, \lambda_i)}\right]^{-\frac{1}{a^*}}, \quad a^* \neq 0.$$

- **Step 5:** To compute the CI of $\theta = g(\alpha, \lambda)$, we order $\theta_1, \theta_2, ..., \theta_M$ as $\theta_{(1)}, \theta_{(2)}, ..., \theta_{(M)}$. Then $(1 - \beta)100\%$ symmetric CI of θ is given by $\left(\theta_{([\frac{M\beta}{2}])}, \theta_{([\frac{M(1-\beta)}{2}])}\right)$.