REVSTAT – Statistical Journal
Volume 17, Number 3, July 2019, 401–427

PREDICTION INTERVALS OF THE RECORD-VALUES PROCESS

Authors: Amany E. Aly
– Department of Mathematics, Faculty of Science, Helwan University, Ain Helwan, Cairo, Egypt
aabusabh@hotmail.com

H.M. Barakat
– Department of Mathematics, Faculty of Science, Zagazig University, Zagazig, Egypt
hbarakat2@hotmail.com

Magdy E. El-Adll
– Department of Mathematics, Faculty of Science, Helwan University, Ain Helwan, Cairo, Egypt
meladll2@hotmail.com

Received: March 2017 Revised: May 2017 Accepted: June 2017

Abstract:

- In this paper, exact prediction intervals of the record-values process are constructed. The record-values process model, may be considered as the collection of record-values with integer or non-integer indices. It includes both usual k-th record-values and fractional k-th record-values models. For constructing the prediction intervals, two predictive pivotal quantities are developed. The distributions of the predictive pivotal quantities are derived and it is revealed that the distribution functions of the predictive pivotal quantities are similar for the upper and lower fractional record-values. More results are obtained for the exponential upper record-values process, including two point predictors and their exact mean square errors. Some efficient algorithms are given and Monte Carlo simulation studies are conducted for comparing pivotal quantities. Finally, three real data sets are analyzed.

Key-Words:

- *record-values process; pivotal quantity; prediction interval; coverage probability; Monte Carlo simulation.*

AMS Subject Classification:

- 62E15, 62G30, 62G32, 62M20, 62F25.

1. INTRODUCTION

Record values arise naturally in many practical problems and there are several situations pertaining to meteorology, hydrology, sporting and athletic events where only record-values may be recorded. Outcomes of competitions, e.g. in athletics, arise in ascending order. In particular, sport events attract many spectators since records and best results appeal to people. Being most popular in sports, lists of best results and records are of particular interest in many other areas of real life as well. For an elaborate treatment on records and their applications see: Arnold *et al.* [6], Nevzorov [40], Gulati and Padget [27], and Ahsanullah ([1], [3]). The first result for record-values involving independent and identically observations was reported by Chandler [17]. Dziubdziela and Kopocinški [21] generalized the concept of record-values of [17] to a more generalized nature and called them k-th record-values. Since the k-th member of the sequence of the classical record-values is also known as the k-th record-value, the record-values defined in [21] is also called generalized record-values. Some properties and applications for current records are given in Barakat *et al.* [11]. Stigler [46] introduced the concept of order statistics process, which may be considered as fractional order statistics for non-integer index. Jones [31] gave an alternative construction of Stigler's uniform fractional order statistics. Namely, ordinary order statistics of a sample from uniform distribution are used to construct random variables (rv's) with the same joint distribution as Stigler's order statistics. Some applications of fractional order statistics are given in Hutson [30]. Bieniek and Szynal [14] follows a similar method of fractional order statistics to introduce the fractional record-values or the record-values process, which can be considered as a family of k-th record-values with n replaced by a positive number t.

One of the most important problems in statistics, is to predict future events based on past or current events. A predictor may be either a point or an interval predictor. Point predictor of future records was studied by Kaminsky and Nelson [32], Ahsanullah [2], Nagaraja [37], and Doganakso and Balakrishnan [19]. Prediction intervals of future records were given in Dunsmore [20], Balakrishnan *et al.* [7], Berred [13], AL-Hussaini and Ahmad [4], and Raqab and Balakrishnan [42]. Bayesian and non-Bayesian approaches have been extensively studied by many authors, e.g. Lawless [36], Kaminsky and Rhodin [34], Geisser [25], Nagaraja [39], and Kaminsky and Nelson [33]. Recent works of prediction using pivotal quantities include, Barakat *et al.* ([8], [9], [10], [12]), El-Adll [22], Aly [5], and El-Adll and Aly [23].

1.1. Motivation of the study

According to Theorem 6.3.1 page 339 of Galambos [24], the ordinary record-values are very rare to be observed. For example, one may wait too many years before observing the next upper (lower) record of the amount of water added to a given river. Although the fractional k-th record-values cannot be observed in practice, prediction of future fractional k-th record-values is prominent in applications. As an example, the prediction of fractional k-th record-values can be applied in reliability and survival data analysis since the fractional k-th record-values may be considered as an estimator of the inverse cumulative hazard function and therefore the quantiles of the population cdf. On the other hand, the employment of the k-th fractional record-values provides an interval estimate with accurate significant level, while the use of the k-th ordinary record-values gives an interval estimate with approximate significant level (e.g., [14]).

Furthermore, our study is carried in a general framework which include prediction of the usual record-values, as well as k-th ordinary record-values, as special cases. Thereby, all the obtained new results not only have theoretical importance but also have practical importance. Thus the results of this paper, which are given in the present general framework, are beneficial when it is necessary to predict the quantiles of a distribution for which, the type of the hazard function may be changed in future.

In the next section, we give a comprehensive survey for the main results of the record-values process, that will be needed in this paper, most of these results are due to [14].

In Sections 3 and 4 of this paper, two prediction intervals for future fractional upper (lower) records are constructed based on two general predictive pivotal quantities. More details for the exponential distribution including, two point predictors and their exact mean square errors are considered for the upper record-value process. In Section 5, two simulation studies are carried out to explain the efficiency of the proposed results. In one of them, the distribution parameters are assumed to be unknown. Some applications to real data are given in Section 6. Two basic algorithms for generation ordinary record-values and fractional upper record-values, as well as an algorithm to implement these prediction intervals, are given in an Appendix.

2. PRELIMINARY RESULTS

In this section, some important preliminary and auxiliary results for the basic distribution theory of ordinary and fractional records are presented. Let $\{X_n,\ n \geqslant 1\}$ be a sequence of independent and identically distributed (iid) random variables (rv's) having a continuous cumulative distribution function (cdf) $F(x)$ and probability density function (pdf) $f(x)$. Furthermore, suppose that $X_{1:n}, X_{2:n}, ..., X_{n:n}$ denote the order statistics of the random sample $X_1, X_2, ..., X_n$.

2.1. Ordinary records

An observation X_j is called an upper record-value if its value exceeds that of all previous observations. Thus, X_j is an upper record-value if $X_j > X_i$ for every $i < j$. In other words, the upper record-value R_n of a random sample of size n can be expressed as $R_n = X_{n:n} = \max\{X_1, X_2, ..., X_n\}$. Dziubdziela and Kopocinški [21] extended the concept of upper record to the k-th upper record, for $k \geqslant 1$, which is formulated in the following definition.

Definition 2.1 (cf. Dziubdziela and Kopocinški [21]). The k-th upper record times, $T_k(n),\ n \geqslant 1$, of the sequence $\{X_i,\ i \geqslant 1\}$ is defined for fixed $k \geqslant 1$, as $T_k(1) = 1$ and

$$T_k(n+1) = \min\{j > T_k(n) : X_{j:j+k-1} > X_{T_k(n):T_k(n)+k-1}\},\ n > 1,$$

and the k-th upper record-values as $R_n^{(k)} = X_{T_k(n):T_k(n)+k-1},\ n \geqslant 1$.

The k-th lower record-values is defined similarly. Clearly, $R_1^{(k)} = X_{1:k} = \min\{X_1, ..., X_k\}$. For $k = 1$ we have $R_n^{(1)} = R_n = X_{n:n}$. In other words, the k-th upper record-sequence is the sequence of the k-th largest yet seen. Although the term "record times"is used in all definitions related to records in statistical literature, it does not mean the time in its verbal sense. The pdf of the k-th upper record-value is

$$f_{R_n^{(k)}}(r) = \frac{k^n}{\Gamma(n)}[H(r)]^{n-1}[\bar{F}(r)]^{k-1}f(r), \quad -\infty < r < \infty, \tag{2.1}$$

where $H(r) = -\log[1 - F(r)]$ is the cumulative hazard function, $h(r) = H'(r) = f(r)/\bar{F}(r)$ denotes the hazard (failure rate) function and $\bar{F} = 1 - F$. The joint pdf of $R_m^{(k)}$ and $R_n^{(k)}$, $m < n$ for $-\infty < r_m < r_n < \infty$, can be written in the form

$$\begin{aligned} &f_{R_m^{(k)},R_n^{(k)}}(r_m, r_n) \\ &= \frac{k^n}{\Gamma(m)\Gamma(n-m)}\left[H(r_m)\right]^{m-1}\left[H(r_n) - H(r_m)\right]^{n-m-1}\left[\bar{F}(r_n)\right]^k h(r_m)h(r_n). \end{aligned} \tag{2.2}$$

Furthermore, the joint pdf of the random vector $(R_1^{(k)}, R_2^{(k)}, ..., R_n^{(k)})$ is given by

$$f_{R_1^{(k)},R_2^{(k)}...,R_n^{(k)}}(r_1, r_2..., r_n) = k^n[\bar{F}(r_n)]^k \prod_{i=1}^{n} h(r_i), \quad -\infty < r_1 < r_2 < ... < r_n < \infty. \tag{2.3}$$

For more details of the previous three relations, see [21], [26], and [38].

2.2. Upper record-values process

Let $\left\{W_n^{(k)},\ n \geqslant 1\right\}$ denote the k-th upper record-values from the standard exponential distribution (EXP(1)). The following two facts, which are due to Ahsanullah [1], characterize the exponential distribution.

Fact 1. For any positive integers m and n, with $m < n$, the rv's $W_m^{(k)}$ and $W_n^{(k)} - W_m^{(k)}$ are independent.

Fact 2. The spacings $W_n^{(k)} - W_m^{(k)}$ follow gamma distribution with parameters $n - m$ and k, respectively.

The following definition, which is due to [14], is necessary to construct the record-value process and fractional k-th record-values.

Definition 2.2. Let $k \in \mathbb{N}$ be fixed and $W^{(k)} = \left\{W^{(k)}(t),\ t \geqslant 0\right\}$ be a stochastic process such that:

(i) $W^{(k)}(0) = 0$ almost sure;

(ii) $W^{(k)}(t)$ has independent increments;

(iii) For every $t > s \geqslant 0$, $W^{(k)}(t) - W^{(k)}(s)$ has gamma distribution with parameters $t - s$ and k, respectively.

Then $\left\{W^{(k)}(t),\ t \geqslant 0\right\}$ is called the exponential k-th upper record-values process. Moreover, the rv's, $W^{(k)}(t)$, $t > 0$, are said to be exponential fractional k-th upper record-values.

Remark 2.1.

1. By fractional k-th record-values, we mean k-th record-values with fractional indices.
2. We shall assume that the cdf F is continuous with pdf f and quantile function
$$F^{-1}(q) = \inf\{v : F(v) \geqslant q\}, \quad 0 \leqslant q < 1.$$

The k-th record-values process and the fractional k-th record-values based on F are formulated in the following definition:

Definition 2.3 (Bieniek and Szynal [14])**.** The stochastic process $Y^{(k)} = \left\{Y^{(k)}(t),\ t \geqslant 0\right\}$, where
$$Y^{(k)}(t) = F^{-1}(1 - \exp[-W^{(k)}(t)]), \quad t \geqslant 0,$$
is called the k-th upper record-values process based on F and the rv's $Y^{(k)}(t)$, $t > 0$, are said to be fractional k-th upper record-values from F.

As in the ordinary record-values, the pdf $f_{Y^{(k)}(t)}(y)$ of the fractional k-th upper record-value $Y^{(k)}(t)$ is

$$(2.4)\qquad f_{Y^{(k)}(t)}(y) = \frac{k^t}{\Gamma(t)}[H(y)]^{t-1}[\bar{F}(y)]^{k-1}f(y), \quad -\infty < y < \infty, \quad t > 0,$$

and the joint pdf of $Y^{(k)}(t_r)$ and $Y^{(k)}(t_s),\ t_s > t_r \geqslant 0$, can be written for $-\infty < y_r < y_s < \infty$, as

$$(2.5)\qquad f_{Y^{(k)}(t_r),Y^{(k)}(t_s)}(y_r, y_s) = \frac{k^{t_s}}{\Gamma(t_r)\Gamma(t_s - t_r)}\,[H(y_r)]^{t_r-1}\,[H(y_s) - H(y_r)]^{t_s-t_r-1}\,[\bar{F}(y_s)]^k h(y_r)h(y_s).$$

Moreover, if $0 = t_0 < t_1 < ... < t_n$, then the joint pdf of the random vector $\mathbf{Y} = (Y^{(k)}(t_1), Y^{(k)}(t_2), ..., Y^{(k)}(t_n))$ is given by (c.f. [14])

$$(2.6)\qquad f_{\mathbf{Y}}(y_{t_1}, y_{t_2}, ..., y_{t_n}) = k^{t_n}[\bar{F}(y_{t_n})]^k \prod_{i=1}^{n} \frac{(H(y_{t_i}) - H(y_{t_{i-1}}))^{t_i - t_{i-1}-1}h(y_{t_i})}{\Gamma(t_i - t_{i-1})},$$

for $-\infty < y_{t_1} < y_{t_2} < ... < y_{t_n} < \infty$.

2.3. Lower record-values process

Let $\left\{Z_n^{(k)},\ n \geqslant 1\right\}$ denote the k-th lower record-values from the standard negative exponential distribution (NEXP(1)), with cdf $G^*(x) = e^x,\ x \leqslant 0$.

Definition 2.4. Let $k \in \mathbb{N}$ be fixed and $Z^{(k)} = \{Z^{(k)}(t),\ t \geqslant 0\}$ be a stochastic process such that:

- **(i)** $Z^{(k)}(0) = 0$ almost sure;
- **(ii)** $Z^{(k)}(t)$ has independent increments;
- **(iii)** For any $t > s \geqslant 0,\ Z^{(k)}(t) - Z^{(k)}(s)$ has a reverse gamma distribution with parameters $t - s$ and k (the reverse gamma pdf is $f(x) = \frac{k^{s-t}}{\Gamma(t-s)}|x|^{t-s-1}e^{x/k},\ x \leq 0$).

Then $\{Z^{(k)}(t),\ t \geqslant 0\}$ is called the negative exponential k-th lower record-values process. Moreover, the rv's $Z^{(k)}(t), t > 0$, are said to be negative exponential fractional k-th lower record-values.

Definition 2.5 (Bieniek and Szynal [14])**.** The stochastic process $X^{(k)} = \left\{X^{(k)}(t),\ t \geqslant 0\right\}$, where

$$X^{(k)}(t) = F^{-1}(\exp[Z^{(k)}(t)]), \quad t \geqslant 0,$$

is called the k-th lower record-values process based on the cdf F and the rv's $X^{(k)}(t), t > 0$, are said to be fractional k-th lower record-values from F.

The pdf $f_{X^{(k)}(t)}(x)$ of the fractional k-th lower record-values $X^{(k)}(t)$ is

$$f_{X^{(k)}(t)}(x) = \frac{k^t}{\Gamma(t)}[-\log F(x)]^{t-1}[F(x)]^{k-1}f(x), \quad -\infty < x < \infty, \quad t > 0, \tag{2.7}$$

and the joint pdf of $X^{(k)}(t_r)$ and $X^{(k)}(t_s)$, $t_s > t_r \geqslant 0$ can be written for $-\infty < x_s < x_r < \infty$, as

$$\begin{aligned} & f_{X^{(k)}(t_r),X^{(k)}(t_s)}(x_r, x_s) \\ &= \frac{k^{t_s}}{\Gamma(t_r)\Gamma(t_s - t_r)}[-\log F(x_r)]^{t_r-1}\left[\log\frac{F(x_r)}{F(x_s)}\right]^{t_s-t_r-1}[F(x_s)]^k\frac{f(x_r)}{F(x_r)}\frac{f(x_s)}{F(x_s)}. \end{aligned} \tag{2.8}$$

3. PREDICTION OF FUTURE UPPER RECORD-VALUES PROCESS

In this section, two predictive pivotal quantities (the pivotal quantity is a function of the sample $X_1, X_2 ...$, and on the distribution parameters, but its distribution does not depend on the distribution parameters) are developed to construct prediction intervals of future fractional upper record-values from a continuous distribution. The following theorem is formulated for the first pivotal quantity, which enables us to predict any future fractional k-th upper record-value $Y^{(k)}(t_s)$ based on one fractional k-th upper record-value $Y^{(k)}(t_r)$ with $r < s$.

Theorem 3.1. *Let $0 = t_0 < t_1 < t_2 < ... < t_n$ be positive real numbers and $Y^{(k)}(t_r)$ be the r-th fractional k-th upper record-values from a continuous distribution with cdf F and pdf f. Then the pdf and the cdf of the pivotal quantity $P_1 = \left(W^{(k)}(t_s) - W^{(k)}(t_r)\right)/W^{(k)}(t_r)$, with $s > r$, respectively, are*

$$f_{P_1}(p_1) = \frac{1}{B(t_s - t_r, t_r)}p_1^{t_s-t_r-1}(1 + p_1)^{-t_s}, \quad p_1 > 0, \tag{3.1}$$

and

$$F_{P_1}(p_1) = I_{\frac{p_1}{1+p_1}}(t_s - t_r, t_r), \quad p_1 \geqslant 0, \tag{3.2}$$

where $I_z(a, b) = \frac{1}{B(a,b)}\int_0^z u^{a-1}(1 - u)^{b-1}du$, $0 < z < 1$, *is the incomplete beta function,* $B(a, b) = \int_0^1 u^{a-1}(1 - u)^{b-1}du$ *and*

$$W^{(k)}(t_i) = -\log\bar{F}(Y^{(k)}(t_i)), \qquad i = 1, 2, ..., n. \tag{3.3}$$

A $100(1 - \delta)\%$ *predictive confidence interval (PCI) for the future fractional k-th upper record-value $Y^{(k)}(t_s)$, (L, U_{P_1}), is*

$$L = Y^{(k)}(t_r) \quad \text{and} \quad U_{P_1} = F^{-1}\left(1 - \left(\bar{F}(Y^{(k)}(t_r))\right)^{1+p_1(\delta)}\right),$$

where $p_1(\delta)$ *can be obtained by solving the non linear equation* $F_{P_1}(p_1(\delta)) = 1-\delta$*. Moreover, when* $Y^{(k)}(t_i) = W^{(k)}(t_i), i = 1, 2, ..., n$ *(i.e.,* F *is EXP(1)), the expected interval width of the PCI is* $\frac{p_1(\delta)t_r}{k}$*. Furthermore,* $\tilde{W}^{(k)}(t_s) = \frac{t_s}{t_r} W^{(k)}(t_r)$ *is an unbiased point predictor based on* P_1*.*

Proof: Since the transformation $w = -\log \bar{F}(y)$ is one to one and onto (monotone increasing function), the pdf of the rv $W^{(k)}(t)$, $f_{W^{(k)}(t)}(w)$, is given by

$$f_{W^{(k)}(t)}(w) \;=\; |J| f_{Y^{(k)}(t)}(y(w)), \quad \text{where} \quad |J| = \frac{dy}{dw} = \frac{1}{f(y)} e^{-w}.$$

Therefore, by (2.4) we have $f_{W^{(k)}(t)}(w) = \frac{k^t}{\Gamma(t)} w^{t-1} e^{-kw}$, $w > 0$, which is the pdf of fractional k-th upper record-values based on EXP(1). Thus, the joint pdf of $W^{(k)}(t_r)$ and $W^{(k)}(t_s)$, $t_s > t_r \geqslant 0$ based on EXP(1) can be written by (2.5) as

$$\text{(3.4)} \qquad f_{W^{(k)}(t_r), W^{(k)}(t_s)}(w_r, w_s) = \frac{k^{t_s}}{\Gamma(t_r)\Gamma(t_s - t_r)} w_r^{t_r - 1} (w_s - w_r)^{t_s - t_r - 1} e^{-k w_s},$$

with $0 < w_r < w_s < \infty$. By a standard method of transformation of rvs, the joint pdf $f_{P_1, W_r^{(k)}}(p_1, w_r)$ of P_1 and $W^{(k)}(t_r)$ can be written in the form

$$\text{(3.5)} \quad f_{P_1, W_r^{(k)}}(p_1, w_r) = \frac{k^{t_s}}{\Gamma(t_r)\Gamma(t_s - t_r)} w_r^{t_s - 1} p_1^{t_s - t_r - 1} e^{-k(1+p_1)w_r}, \quad p_1 > 0, \;\; w_r > 0.$$

Thus, we have

$$\begin{aligned} f_{P_1}(p_1) \;&=\; \int_0^\infty f_{P_1, W_r^{(k)}}(p_1, w_r) dw_r \\ &=\; \frac{k^{t_s}}{\Gamma(t_r)\Gamma(t_s - t_r)} \int_0^\infty w_r^{t_s - 1} p_1^{t_s - t_r - 1} e^{-k(1+p_1)w_r} dw_r. \end{aligned}$$

By the definition of gamma function, the above integration can be simplified in the form (3.1). Moreover, the cdf of the pivotal quantity P_1 is given by

$$\begin{aligned} F_{P_1}(p_1) \;&=\; \int_0^{p_1} f_{P_1}(z) dz = \int_0^{p_1} \frac{1}{B(t_r, t_s - t_r)} \, z^{t_s - t_r - 1} (1+z)^{-t_s} dz \\ &=\; \frac{1}{B(t_s - t_r, t_r)} \int_0^{p_1} \left(\frac{z}{1+z}\right)^{t_s} \left(\frac{1}{z}\right)^{t_r + 1} dz. \end{aligned}$$

If we set $w = \dfrac{z}{1+z}$ in the above integration, it yields (3.2). If δ is such that $F_{P_1}(p_1(\delta)) = P(P_1 \leqslant p_1(\delta)) = 1 - \delta$, we can write

$$\begin{aligned} 1 - \delta \;&=\; P\left(0 < \frac{W^{(k)}(t_s) - W^{(k)}(t_r)}{W^{(k)}(t_r)} \leqslant p_\delta\right) = P\left(0 < \frac{W^{(k)}(t_s)}{W^{(k)}(t_r)} - 1 \leqslant p_1(\delta)\right) \\ &=\; P\left(W^{(k)}(t_r) < W^{(k)}(t_s) \leqslant (1 + p_1(\delta))\, W^{(k)}(t_r)\right) = P\left(L < Y^{(k)}(t_s) \leqslant U_{P_1}\right). \end{aligned}$$

The expected interval width of the PCI for $W^{(k)}(t_s)$ is given by

$$E\left[(1+p_1(\delta))W^{(k)}(t_r) - W^{(k)}(t_r)\right] = E\left[p_1(\delta)W^{(k)}(t_r)\right] = \frac{p_1(\delta)t_r}{k}.$$

Finally, a point predictor based on P_1 can be obtained from the relation $\tilde{W}^{(k)}(t_s) = L + c_1(U_{P_1} - L)$, where the constant c_1 is such that $E[\tilde{W}^{(k)}(t_s)] = E[W^{(k)}(t_s)] = t_s/k$. Hence the theorem. □

Theorem 3.2. *Assume that* $0 = t_0 < t_1 < t_2 < ... < t_n$ *are positive real numbers. Furthermore, let* $Y^{(k)}(t_1)$ *and* $Y^{(k)}(t_r)$ *be the first and the* r*-th fractional* k*-th upper record-values from a continuous distribution with cdf* F *and pdf* f*. Then the pdf and the cdf of the pivotal quantity*

$$P_2 = \frac{W^{(k)}(t_s) - W^{(k)}(t_r)}{W^{(k)}(t_r) - W^{(k)}(t_1)}, \qquad s > r > 1, \tag{3.6}$$

are given by

$$f_{P_2}(p_2) = \frac{1}{B(t_s - t_r, t_r - t_1)}(1+p_2)^{-(t_s - t_1)} p_2^{t_s - t_r - 1}, \quad p_2 > 0, \tag{3.7}$$

and

$$F_{P_2}(p_2) = I_{\frac{p_2}{1+p_2}}(t_s - t_r, t_r - t_1), \quad p_2 \geqslant 0, \tag{3.8}$$

respectively, with $W^{(k)}(t_i) = -\log \bar{F}(Y^{(k)}(t_i)), i = 1, 2, ..., n$. *A* $100(1-\delta)\%$ *PCI for the future* k*-th upper record-value* $Y^{(k)}(t_s)$ *is* (L, U_{P_2})*, with*

$$U_{P_2} = F^{-1}\left(1 - \bar{F}(Y^{(k)}(t_1))\left(\frac{\bar{F}(Y^{(k)}(t_r))}{\bar{F}(Y^{(k)}(t_1))}\right)^{1+p_2(\delta)}\right),$$

where $p_2(\delta)$ *can be obtained by solving the non linear equation* $F_{P_2}(p_2(\delta)) = 1 - \delta$. *Moreover, an unbiased point predictor based on* P_2 *is given by*

$$\hat{W}^{(k)}(t_s) = W^{(k)}(t_r) + \left(\frac{t_s - t_r}{t_r - t_1}\right)\left(W^{(k)}(t_r) - W^{(k)}(t_1)\right), \quad s > r > 1,$$

which is the best linear unbiased predictor (BLUP) for $W^{(k)}(t_s)$*.*

Proof: We see from the proof of Theorem 3.1 that the rv $W^{(k)}(t_i)$, $i = 1, 2, ..., n$, can be expressed as fractional k-th upper record-values based on EXP(1). Therefore, the joint pdf of $W^{(k)}(t_1)$, $W^{(k)}(t_r)$ and $W^{(k)}(t_s)$ is given by

$$f_{1,r,s}(w_1, w_r, w_s) = \frac{k^{t_s}}{\Gamma(t_1)\Gamma(t_r - t_1)\Gamma(t_s - t_r)} w_1^{t_1 - 1}(w_r - w_1)^{t_r - t_1 - 1}(w_s - w_r)^{t_s - t_r - 1} e^{-k w_s},$$

for $0 < w_1 < w_r < w_s < \infty$, where for simplicity we write W_i instead of $W^{(k)}(t_i)$. On the other hand, by using the linear transformations $U = W_1, V = W_r - W_1$ and $W = W_s - W_r$, the joint pdf of the rv's $U,\ V$ and W is

$$f_{U,V,W}(u,v,w) = \frac{k^{t_s}}{\Gamma(t_1)\Gamma(t_r - t_1)\Gamma(t_s - t_r)} u^{t_1-1} v^{t_r-t_1-1} w^{t_s-t_r-1} \exp(-k(u+v+w)), \tag{3.9}$$

with $u > 0, v > 0, w > 0$. The joint pdf of U, V and $P_2 = W/V$ can be written as

$$f_{U,V,P_2}(u,v,p_2) = \frac{k^{t_s}}{\Gamma(t_1)\Gamma(t_r - t_1)\Gamma(t_s - t_r)} u^{t_1-1} v^{t_s-t_1-1}\, p_2^{t_s-t_r-1} \exp\left(-k\left[u + (1+p_2)\,v\right]\right),$$

for $u > 0,\ v > 0,\ p_2 > 0$. Thus, the pdf of the pivotal quantity P_2 is

$$f_{p_2}(p_2) = \int_0^\infty \int_0^\infty f_{U,V,P_2}(u,v,p_2)\,du\,dv.$$

By evaluating the above integration, we get (3.7) and (3.8). Moreover, we have

$$\begin{aligned} 1-\delta &= F_{P_2}(p_2) = P\left(P_2 \leqslant p_2(\delta)\right) = P\left(0 < \frac{W_s - W_r}{W_r - W_1} \leqslant p_2(\delta)\right) \\ &= P\left(W_r < W_s \leqslant W_r + p_2(\delta)(W_r - W_1)\right) = P\left(L < Y^{(k)}(t_s) \leqslant U_{P_2}\right). \end{aligned}$$

Furthermore, the expected interval width for the PCI of $W^{(k)}(t_s)$ is given by

$$E\left[p_2(\delta)(W_r - W_1)\right] = \frac{p_2(\delta)}{k}(t_r - t_1).$$

Finally, we can obtain the point predictor, $\hat{W}^{(k)}(t_s)$, as in Theorem 3.1. By the same method of [2] (with a suitable modifications), it is not difficult to verify that $\hat{W}^{(k)}(t_s)$ is the BLIP. Hence the theorem. □

4. PREDICTION OF FUTURE LOWER RECORD-VALUES PROCESS

In this section, the predictive pivotal quantities presented in Section 3, will be modified to construct prediction intervals of future fractional lower record-values from continuous distributions.

Theorem 4.1. *Let* $0 = t_0 < t_1 < t_2 < ... < t_n$ *be positive real numbers and* $X^{(k)}(t_1), X^{(k)}(t_2), ...,\ X^{(k)}(t_r)$ *be the first* r *fractional* k*-th lower record-values from a continuous distribution whose pdf* f *and cdf* F*. Then the pdf and* *the cdf of the pivotal quantity* $P_1^* = \left(Z^{(k)}(t_s) - Z^{(k)}(t_r)\right)/Z^{(k)}(t_r)$ *are given by (3.1) and (3.2), respectively, where*

$$Z^{(k)}(t_i) = \log F(X^{(k)}(t_i)), \qquad i = 1, 2, ..., n. \tag{4.1}$$

A $100(1-\delta)\%$ *PCI for the future fractional k-th lower record-value* $X^{(k)}(t_s)$ *is* $(L_{P_1^*}, U)$ *where*

$$L_{P_1^*} = F^{-1}\left(\left(F(X^{(k)}(t_r)\right)^{1+p_1^*(\delta)}\right), \qquad U = X^{(k)}(t_r),$$

and p_1^* *can be obtained by solving the non linear equation* $F_{P_1^*}(p_1^*) = 1-\delta$*. Moreover, when* $X^{(k)}(t_i) = Z^{(k)}(t_i), i = 1, 2, ..., n$ *(i.e., F is* NEXP(1)*), the expected interval width of the PCI is* $\frac{p_1^*(\delta)t_r}{k}$*.*

Proof: Since the transformation $Z = \log F(y)$ is one to one and onto (monotone increasing function), the pdf $f_{Z^{(k)}(t)}(z)$ of the rv $Z^{(k)}(t)$ is given by $f_{Z^{(k)}(t)}(z) = |J| f_{X^{(k)}(t)}(x(z))$, where $|J| = \frac{dx}{dz} = \frac{1}{f(x)}e^z$. Therefore, by (2.7) we have $f_{Z^{(k)}(t)}(z) = \frac{k^t}{\Gamma(t)}(-z)^{t-1}e^{kz}$, $z < 0$, which is the pdf of fractional k-th lower record-values based on NEXP(1). The remaining part of the proof is similar to the corresponding part of the proof of Theorem 3.1, with only obvious changes. □

Theorem 4.2. *Let* $0 = t_0 < t_1 < t_2 < ... < t_n$ *be positive real numbers and* $X^{(k)}(t_1)$ *and* $X^{(k)}(t_r)$ *be the first and the r-th fractional k-th lower record-values from a continuous distribution whose pdf f and cdf F. Then the pdf and the cdf of the pivotal quantity*

$$P_2^* = \frac{Z^{(k)}(t_r) - Z^{(k)}(t_s)}{Z^{(k)}(t_1) - Z^{(k)}(t_r)}, \; s > r > 1, \tag{4.2}$$

are given by (3.7) and (3.8) respectively, with $Z^{(k)}(t_i) = \log F(X^{(k)}(t_i)),\; i = 1, 2, ..., n$*. A* $100(1-\delta)\%$ *PCI for the future k-th lower record-value* $X^{(k)}(t_s)$ *is* $(L_{P_2^*}, U)$ *where*

$$L_{P_2^*} = F^{-1}\left(F(X^{(k)}(t_r))\left(\frac{F(X^{(k)}(t_r))}{F(X^{(k)}(t_1))}\right)^{p_2^*(\delta)}\right), \qquad U = X^{(k)}(t_r),$$

and $p_2^*(\delta)$ *can be obtained by solving the non linear equation* $F_{P_2^*}(p_2^*(\delta)) = 1-\delta$*.*

Proof: The joint pdf of the fractional k-th lower record-values $Z^{(k)}(t_1)$, $Z^{(k)}(t_r)$ and $Z^{(k)}(t_s)$ based on NEXP(1) is given by

$$f_{1,r,s}(z_1, z_r, z_s) \\ = \frac{k^{t_s}}{\Gamma(t_1)\Gamma(t_r - t_1)\Gamma(t_s - t_r)}(-z_1)^{t_1-1}(z_1 - z_r)^{t_r-t_1-1}(z_r - z_s)^{t_s-t_r-1}e^{kz_s},$$

$-\infty < z_s < z_r < z_1 \leqslant 0$. Now, consider the linear transformations $U^* = -Z_1$, $V^* = Z_1 - Z_r$ and $W^* = Z_r - Z_s$, the joint pdf of the rv's U^*, V^* and W^* is given by relation (3.9). Therefore, the rest of the proof is similar, with only obvious changes, to the corresponding part of the proof of Theorem 3.2. □

Remark 4.1.

1. The preceding results can be proved by the independence between the components in each of the vectors $(W^{(k)}(t_r), W^{(k)}(t_s) - W^{(k)}(t_r))$, $(W^{(k)}(t_1), W^{(k)}(t_r) - W^{(k)}(t_1), W^{(k)}(t_s) - W^{(k)}(t_r))$, $(Z^{(k)}(t_r),\ Z^{(k)}(t_s) - Z^{(k)}(t_r))$ and $(Z^{(k)}(t_1), Z^{(k)}(t_r) - Z^{(k)}(t_1), Z^{(k)}(t_s) - Z^{(k)}(t_r))$.

2. The lower and the upper limits of the PCI for future fractional k-th upper (lower) record-values depend on the population cdf F.

3. The ordinary upper (lower) record-values are obtained as special cases from the presented methods by setting $t_i = i$, for all $i = 1, 2, ..., n$.

4. All the preceding results remain valid if we replace $t_i = i,\ i = 1, 2, ..., r$, that is, fractional k-th upper (lower) record-values can be predicted via ordinary k-th upper (lower) record-values.

5. SIMULATION STUDIES

In this section, simulation studies are conducted to demonstrate the efficiency of the presented results. For this purpose, three algorithms are established in Appendix A.

Let us first check the validity of the first two algorithms, by generating ten fractional upper records $Y^{(1)}(t_i), i = 1, 2, ..., 10$, (see Table 1) based on Weibull distribution with shape and scale parameters $\alpha = 3$ and $\beta = 30$, respectively. It is easy to compute the theoretical expectation of each of these records, namely,

$$E[Y^{(k)}(t_i)] = \beta k^{-\frac{1}{\alpha}} \frac{\Gamma\left(t_i + 1/\alpha\right)}{\Gamma(t_i)}, \quad i = 1, 2, ..., 10. \tag{5.1}$$

The idea of this simple test is to compare the theoretical value $E[Y^{(k)}(t_i)]$ with the estimated value resulted from application of the algorithms, i.e., the average value $\bar{Y}^{(k)}(t_i)$. In order to compute the average value of each of these records, we repeat the generation processes of these ten records, different values of times, $M = 10^3,\ 10^4,\ 10^5,\ 10^6$, and for each of these replicates M, we compute the average $\bar{Y}^{(k)}(t_i)$, for each i. Table 1 summarizes these computations and shows that the theoretical expectations for all records are close to the estimated values, which are resulted via the application of the two algorithms.

All Computations are performed by using Mathematica version 10 with processor: Intel(R) Core(TM) i7-2640 cpu @ 2.80GHz 2.80GHz, RAM 4.00GB, and system type 64-bit operating system.

Table 1: A comparison between $\bar{Y}^{(1)}(t_i)$ and $E[Y^{(1)}(t_i)]$.

i	1	2	3	4	5	6	7	8	9	10	Run Time(s)
t_i	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	
$\bar{Y}^{(1)}(t_i)$, $M=10^3$	26.7135	31.7547	35.6315	38.8247	41.5164	43.8532	46.0244	48.2064	50.0735	51.8080	0.515
$\bar{Y}^{(1)}(t_i)$, $M=10^4$	26.8238	31.8600	35.7033	38.8794	41.6359	44.0795	46.3186	48.3856	50.2302	51.9270	2.840
$\bar{Y}^{(1)}(t_i)$, $M=10^5$	26.7416	31.8409	35.7152	38.9153	41.6562	44.1004	46.2950	48.3030	50.1555	51.8929	25.896
$\bar{Y}^{(1)}(t_i)$, $M=10^6$	26.7875	31.8445	35.7116	38.9043	41.6630	44.0974	46.2920	48.2962	50.1461	51.8737	267.182
$E[Y^{(1)}(t_i)]$	26.7894	31.8425	35.7192	38.9186	41.6724	44.1078	46.3026	48.3085	50.1612	51.8869	0.265

5.1. Exact and numerical computations

The exact expected values of the upper limits for the future fractional k-th upper record-value, $Y^{(k)}(t_s)$, from the exponential distribution with mean $1/\lambda$, based on the pivotal quantities P_1 and P_2, respectively, are given by

$$E[U_{P_1}] = \frac{(1+p_1(\delta))t_r}{\lambda k} \quad \text{and} \quad E[U_{P_2}] = \frac{1}{\lambda k}\left[t_r + p_2(\delta)(t_r - t_1)\right].$$

Moreover, the exact mean square errors of U_{P_1} and U_{P_2}, respectively, are given by

$$\begin{aligned} MSE_{U_{P_1}} &= E\left[U_{P_1} - Y^{(k)}(t_{s+1})\right]^2 \\ &= \frac{1}{(\lambda k)^2}\left[(1+p_1(\delta))^2 t_r(1+t_r) + t_{s+1}(1+t_{s+1}) - 2t_r(1+t_{s+1})(1+p_1(\delta))\right] \end{aligned}$$

and

$$MSE_{U_{P_2}} = E\left[U_{P_2} - Y^{(k)}(t_{s+1})\right]^2$$
$$= \frac{1}{(\lambda k)^2}\left\{p_2(\delta)(t_r - t_1)\left[p_2(\delta)(t_r - t_1 + 1) - 2(t_{s+1} - t_r)\right] + (t_{s+1} - t_r)(t_{s+1} - t_r + 1)\right\}.$$

The mean square predictive errors, based on P_1 and P_2 respectively, are

$$MSE_{P_1} = E\left[\tilde{Y}^{(k)}(t_s) - Y^{(k)}(t_s)\right]^2 = \frac{t_s(t_s - t_r)}{(\lambda k)^2 t_r}, \quad r \geqslant 1,$$

and

$$MSE_{P_2} = E\left[\hat{Y}^{(k)}(t_s) - Y^{(k)}(t_s)\right]^2 = \frac{(t_s - t_r)(t_s - t_1)}{(\lambda k)^2(t_r - t_1)}, \quad s > r > 1,$$

where $\tilde{Y}^{(k)}(t_s)$ and $\hat{Y}^{(k)}(t_s)$ denote the point predictors of $Y^{(k)}(t_s)$ based on the pivotal quantities P_1 and P_2, respectively.

Remark 5.1. Clearly,

$$MSE_{P_2} - MSE_{P_1} = \frac{t_1(t_s - t_r)^2}{(\lambda k)^2 t_r (t_r - t_1)} > 0.$$

That is, $MSE_{P_2} > MSE_{P_1}$, for all $s > r > 1$.

The estimated root mean square errors for the upper limits of the PCI, respectively are defined by

$$R\hat{M}SE_{P_j} = \left[\frac{1}{M-1} \sum_{i=1}^{M} \left(U_{P_j}(i) - Y^{(k)}(t_{i+1}) \right)^2 \right]^{1/2}, \ j = 1, 2.$$

Throughout this paper the following abbreviations are used:

$\bar{Y}^{(k)}(t_s)$: The mean of fractional k-th record-value, $Y^{(k)}(t_s)$, is defined by $\bar{Y}^{(k)}(t_s) = \frac{1}{M} \sum_{i=1}^{M} Y_i^{(k)}(t_s)$, where M denote the number of replicants.

PCI: The predictive confidence interval of future fractional upper record.

$CP_i\%$: The percent of coverage probability based on $P_i,\ i = 1, 2,$ at $\delta = 0.10$.

$(\bar{L}, \bar{U}_{P_i})$: The average lower (upper) limits for the PCI of future fractional upper record.

$(\bar{L}_{P_i^*}, U)$: The average lower (upper) limits for the PCI of future fractional lower record.

$BLUP$: Best linear unbiased predictor for future fractional upper record.

$E[U_{P_i}]$: The expected value of the upper limit of the PCI based on $P_i,\ i = 1, 2$.

$RMSE_{P_i}$: The exact root mean square error for the upper limit of the PCI based on $P_i,\ i = 1, 2$.

$R\hat{M}SE_{P_i}$: The estimated root mean square error for the upper limit of the PCI based on $P_i,\ i = 1, 2$.

The rest of this section contains illustrations of the purposed methods through two simulation studies. The first study for EXP(0.1) is based on $M = 10^5$ replicates of $n = 25$ k-th upper records (including 13 ordinary records and 12 fractional records) corresponding to $t_i = 1, 1.5, 2, 2.5, ..., 13,\ k = 2$. In this study the first $r = 15$ upper records $Y^{(2)}(1), Y^{(2)}(1.5), ..., Y^{(2)}(8)$ are assumed to be known and the next future 9 upper records, $Y^{(2)}(8.5), Y^{(2)}(9), ..., Y^{(2)}(12.5)$ are to be predicted. The results which are shown in Table 2, include 90% coverage probability, two point predictors as well as two prediction intervals and the expected values of the upper limits. Moreover the exact root mean square errors for the point predictors, exact and estimated root mean square errors for the upper limits are given between parentheses. It is worth to mention here that the PCI's as well as the point predictors does not depend on the scale parameter β.

Table 2: Prediction of future ordinary and fractional upper records from EXP(0.1) based on $M = 10^5$ replicates.

t_r	t_s	$CP_1\%$	$CP_2\%$	$\bar{L}$	$\bar{Y}^{(2)}(t_s)$	$\tilde{Y}^{(2)}(t_s)$	$\hat{Y}^{(2)}(t_s)$	$\bar{U}_{P_1}$	$E[U_{P_1}]$	$\bar{U}_{P_2}$	$E[U_{P_2}]$
8.0	8.5	89.914	89.958	40.026	42.525	42.527 (3.644)	42.531 (3.660)	47.651 (6.283)	47.620 (6.255)	47.798 (6.449)	47.756 (6.418)
	9.0	89.998	89.993	40.026	45.031	45.029 (5.303)	45.037 (5.345)	53.375 (9.698)	53.341 (9.688)	53.688 (10.108)	53.632 (10.082)
	9.5	90.064	90.092	40.026	47.527	47.531 (6.673)	47.542 (6.748)	58.501 (12.820)	58.464 (12.816)	58.983 (13.467)	58.917 (13.440)
	10.0	90.014	90.032	40.026	50.037	50.032 (7.906)	50.047 (8.018)	63.368 (15.751)	63.327 (15.740)	64.025 (16.638)	63.947 (16.597)
	10.5	89.989	90.022	40.026	52.531	52.534 (9.057)	52.553 (9.210)	68.091 (18.544)	68.047 (18.535)	68.925 (19.676)	68.837 (19.628)
	11.0	89.967	90.024	40.026	55.040	55.035 (10.155)	55.058 (10.351)	72.722 (21.246)	72.675 (21.242)	73.737 (22.626)	73.638 (22.574)
	11.5	90.036	90.097	40.026	57.541	57.537 (11.215)	57.564 (11.456)	77.290 (23.887)	77.240 (23.887)	78.489 (25.518)	78.380 (25.461)
	12.0	89.998	90.132	40.026	60.055	60.039 (12.247)	60.069 (12.536)	81.812 (26.499)	81.760 (26.484)	83.197 (28.380)	83.078 (28.303)
	12.5	89.983	90.063	40.026	62.550	62.540 (13.258)	62.575 (13.595)	86.301 (29.053)	86.245 (29.046)	87.873 (31.193)	87.744 (31.112)
9.5	10.0	89.993	90.020	47.527	50.037	50.028 (3.627)	50.031 (3.638)	55.006 (6.100)	54.975 (6.083)	55.106 (6.211)	55.066 (6.190)
	10.5	90.003	90.144	47.527	52.531	52.529 (5.257)	52.536 (5.286)	60.562 (9.275)	60.528 (9.270)	60.774 (9.547)	60.723 (9.530)
	11.0	90.041	90.123	47.527	55.040	55.031 (6.589)	55.040 (6.642)	65.514 (12.145)	65.478 (12.152)	65.840 (12.579)	65.781 (12.565)
	11.5	90.034	90.145	47.527	57.541	57.532 (7.780)	57.545 (7.859)	70.202 (14.814)	70.163 (14.827)	70.646 (15.410)	70.577 (15.396)
	12.0	90.069	90.078	47.527	60.055	60.034 (8.885)	60.049 (8.993)	74.741 (17.366)	74.699 (17.370)	75.305 (18.124)	75.228 (18.096)
	12.5	89.965	90.104	47.527	62.550	62.535 (9.934)	62.554 (10.073)	79.185 (19.810)	79.140 (19.820)	79.871 (20.739)	79.786 (20.706)
11.0	11.5	89.975	89.967	55.040	57.541	57.542 (3.615)	57.545 (3.623)	62.417 (5.981)	62.371 (5.964)	62.490 (6.062)	62.437 (6.039)
	12.0	89.942	90.001	55.040	60.055	60.044 (5.222)	60.049 (5.244)	67.856 (8.990)	67.807 (8.976)	68.010 (9.187)	67.946 (9.161)
	12.5	90.016	90.042	55.040	62.550	62.546 (6.528)	62.554 (6.567)	72.687 (11.707)	72.634 (11.685)	72.923 (12.022)	72.851 (11.979)

In the second simulation study we assume that the first five ordinary k-th upper record-values, $Y^{(3)}(1), ..., Y^{(3)}(5)$, have been observed from Weibull distribution with cdf,

$$F(y) = 1 - \exp\left[-\left(\frac{y}{\beta}\right)^{\alpha}\right], \; y > 0, \; \alpha > 0, \; \beta > 0,$$

for $\alpha = 2.5, \; \beta = 40$ and we have to predict the next three ordinary k-th upper record-values and three fractional k-th upper record-values:

$$Y^{(3)}(5.5), \; Y^{(3)}(6), \; Y^{(3)}(6.5), \; Y^{(3)}(7), \; Y^{(3)}(7.5), \; Y^{(3)}(8).$$

The prediction results are obtained in the following two situations:

(**a**) The parameters are assumed to be known.

(**b**) The parameters are unknown and should be estimated.

The maximum likelihood estimators (MLE's) of the parameters based on the first observed $r < n$ ordinary k-th upper record-values can be obtained by maximizing (2.6) (after replacing n with r). Namely,

$$\hat{\alpha} = \frac{r}{\sum_{i=1}^{r-1} \ln(Y^{(k)}(r)/Y^{(k)}(i))} \quad \text{and} \quad \hat{\beta} = \left(\frac{k}{r}\right)^{\frac{1}{\hat{\alpha}}} Y^{(k)}(r). \tag{5.2}$$

But the MLE's are biased and Wang and Ye [47] obtained the corrected unbiased estimators, which are

$$\tilde{\alpha} = \frac{r-2}{\sum_{i=1}^{r-1} \ln(Y^{(k)}(r)/Y^{(k)}(i))} \quad \text{and} \quad \tilde{\beta} = \frac{\Gamma(r)}{\Gamma(r+1/\tilde{\alpha})}\left(1 + \frac{\ln r}{r\tilde{\alpha}}\right)^{r-1} \hat{\beta}. \tag{5.3}$$

Moreover, an unbiased point predictor, $\tilde{Y}^{(k)}(t_s)$ based on P_1 is obtained, and is given by

$$\tilde{Y}^{(k)}(t_s) = \frac{\Gamma(t_r)\Gamma(t_s + 1/\alpha)}{\Gamma(t_s)\Gamma(t_r + 1/\alpha)}\, Y^{(k)}(t_r). \tag{5.4}$$

For each value of t_r and t_s in Table 3, the prediction results obtained based on the exact values of parameters are given in the first two lines, while when the parameters are estimated from (5.3), the prediction results are shown in the last two lines of the same value of t_r and t_s.

Table 3: Prediction of future ordinary and fractional k-th upper record-values with $k = 3$ from Weibull(2.5, 40) based on $M = 10^5$ replicates. The root mean square errors are between parentheses.

$t_r = r$	t_s	$CP_1\%$	$CP_2\%$	L	$\bar{Y}^{(3)}(t_s)$	$\tilde{Y}^{(3)}(t_s)$	U_{P_1}	U_{P_2}	$E[U_{P_1}]$
5.0	5.5	89.969	89.894	47.9290	49.9014	49.9012 (2.7719)	53.6961 (4.4844)	53.9554 (4.7718)	53.6622
		85.578	87.335	47.9290	49.9014	49.8351 (2.9267)	53.6967 (5.4088)	54.2708 (5.9271)	
	6.0	90.045	89.978	47.9290	51.7529	51.7633 (3.8825)	57.6233 (6.5626)	58.1218 (7.1876)	57.5869
		83.096	85.544	47.9290	51.7529	51.6495 (4.3132)	57.6239 (8.5196)	58.7394 (9.6642)	
	6.5	90.026	89.940	47.9290	53.5099	53.5303 (4.7093)	60.8786 (8.2831)	61.5855 (9.1853)	60.8402
		81.141	84.075	47.9290	53.5099	53.3848 (5.4889)	60.8792 (11.2884)	62.5114 (12.9732)	
	7.0	89.959	89.991	47.9290	55.2068	55.2142 (5.4164)	63.7723 (9.7591)	64.6657 (10.9042)	63.7320
		79.384	82.639	47.9290	55.2068	55.0509 (6.5917)	63.7729 (13.8610)	65.9118 (16.0448)	
	7.5	89.954	89.986	47.9290	56.8109	56.8245 (6.0246)	66.4209 (11.0487)	67.4835 (12.4157)	66.3789
		78.043	81.511	47.9290	56.8109	56.6559 (7.6409)	66.4216 (16.2993)	69.0622 (18.9581)	
	8.0	90.013	90.081	47.9290	58.3474	58.3693 (6.5493)	68.8847 (12.1988)	70.1027 (13.7668)	68.8412
		76.884	80.553	47.9290	58.3474	58.2064 (8.6456)	68.8854 (18.6427)	72.0256 (21.7588)	
6.0	6.5	90.137	90.106	51.7529	53.5099	53.5196 (2.4750)	56.8502 (3.9648)	57.0023 (4.1292)	56.8256
		86.360	87.745	51.7529	53.5099	53.4529 (2.5871)	56.8507 (4.5650)	57.1463 (4.8598)	
	7.0	90.007	89.930	51.7529	55.2068	55.2031 (3.4957)	60.3395 (5.7578)	60.6362 (6.1244)	60.3135
		84.132	86.076	51.7529	55.2068	55.0810 (3.8025)	60.3401 (7.0049)	60.9186 (7.6738)	
	7.5	90.001	89.976	51.7529	56.8109	56.8131 (4.2573)	63.2519 (7.2441)	63.6779 (7.7866)	63.2247
		82.368	84.637	51.7529	56.8109	56.6457 (4.8119)	63.2526 (9.1503)	64.1030 (10.1386)	
	8.0	90.016	89.958	51.7529	58.3474	58.3576 (4.8783)	65.8570 (8.5096)	66.4008 (9.2079)	65.8286
		80.930	83.479	51.7529	58.3474	58.1539 (5.7231)	65.8577 (11.1011)	66.9755 (12.3802)	
7.0	7.5	90.086	90.102	55.2068	56.8109	56.8169 (2.2653)	59.8084 (3.5675)	59.9066 (3.6735)	59.7786
		86.856	87.946	55.2068	56.8109	56.7554 (2.3490)	59.8090 (3.9979)	59.9720 (4.1891)	
	8.0	90.164	90.231	55.2068	58.3474	58.3615 (3.1695)	62.9703 (5.1441)	63.1637 (5.3853)	62.9389
		84.952	86.511	55.2068	58.3474	58.2460 (3.4058)	62.9709 (6.0285)	63.2955 (6.4710)	

6. DATA ANALYSIS

In the inverse sampling plan, one takes observations until a fixed number r of records is reached (cf. [28]). According to Hofmann and Nagaraja [28], the amount of Fisher information (FI) for both fixed sample and inverse sampling plan based on all upper records and their record times is greater than the amount of FI based on only upper records. Therefore, in this section, we shall estimate the parameters from the likelihood function for record-breaking (record-values and their inter-record times).

For the inverse sampling plan, the joint likelihood of the upper record-values $Y_1, Y_2, ..., Y_m$ and the inter-record times $\tau_1, \tau_2, ..., \tau_m$ is given by

$$L(\mathbf{y}, \tau; \Theta) = f(y_1, ..., y_m, \tau_1, ..., \tau_m; \Theta) = \prod_{i=1}^{m} f(y_i; \Theta)(F(y_i; \Theta))^{\tau_i - 1}, \tag{6.1}$$

where $\mathbf{y}$ is the vector of observed upper records, $\tau_i, i = 1, 2, ..., m-1,$ are the number of trials following the observation y_i that are needed to obtain the next upper record-value y_{i+1} with $\tau_m = 1$ and Θ is an unknown vector of parameters (e.g. [40], [28] and [35]). Similar result for lower record-breaking is given in [45], [29] and [27], that is

$$L^*(\mathbf{x}, \tau^*; \Theta) = f(x_1, ..., x_m, \tau_1^*, ..., \tau_m^*; \Theta) = \prod_{i=1}^{m} f(x_i; \Theta)(1 - F(x_i; \Theta))^{\tau_i^* - 1}, \tag{6.2}$$

where $\mathbf{x}$ is the vector of observed lower records, $\tau_i^*, i = 1, 2, ..., m-1,$ is the number of trials needed, following x_i to obtain the next lower record x_{i+1}, $\tau_m^* = 1$ and $F(.)$ is cdf of the population from which the sample is drawn. In the rest of this section, three examples to real data are analyzed.

Example 6.1 (Maximum annual temperature)**.** The following data from Long Beach, California, represents the maximum annual temperature in Fahrenheit from 1990 to 2012:

86.7, 81.7, 84.3, 86.4, 84.9, 85.1, 89.7, 82.3, 84.2, 85.8, 81.5, 82.4,
84.3, 84.1, 90.5, 89.4, 87.5, 88.4, 90.3, 84.1, 88.4, 83.0, 86.6.

The upper records and inter-record times for the above data are $y_1 = 86.7,\ y_2 = 89.7,\ y_3 = 90.5$ and $\tau_1 = 6,\ \tau_2 = 8,\ \tau_3 = 1$. First we fit the complete data to some probability distributions. The preliminary fitting indicates that Weibull, extreme value, Frechet distributions are appropriate models for this data. Moreover, the maximum likelihood estimates (MLE's) of parameters are obtained based on (6.1) and then a comparison is performed according to Akaike information criterion (AIC) to select the best model. The results are summarized in Table 4.

Table 4: Comparison between three different distributions via the log Likelihood and AIC.

Model	Parameters	$\mathcal{L} = Log\, L$	AIC
Frechet(α, β)	$\hat{\alpha} = 52.7598,\ \hat{\beta} = 85.2217$	-11.347	26.694
$EVD(\alpha, \beta)$	$\hat{\alpha} = 85.1787,\ \hat{\beta} = 1.67541$	-11.331	26.662
Weibull(α, β)	$\hat{\alpha} = 26.6179,\ \hat{\beta} = 86.4405$	-11.052	26.103

According to AIC Weibull distribution is better than extreme value distribution (EVD) and Frechet distributions. Based on the first three records the upper limits for the next two records and the next two half fractional records are obtained in Table 5.

Table 5: Point predictor and 95% PCI for the next two half record-values and the two record-values for annual maximum temperatures based on the first 3 records.

t_r	t_s	L	$\tilde{Y}^{(1)}(t_s)$	U_{P_1}	$E[U_{P_1}]$	U_{P_2}	$E[U_{P_2}]$
3	3.5	90.500	91.111	92.884	91.873	93.395	91.946
	4	90.500	91.633	93.960	92.936	94.720	93.533
	4.5	90.500	92.089	94.735	93.704	95.652	94.594
	5	90.500	92.494	95.358	94.320	96.385	95.406

Example 6.2 (Maximum annual earthquakes)**.** The data consists of 151 magnitude of the annual maximum earthquakes in the United States during the period from 1769 to 1989 (some data are missing). The data are from Mathematica Documentation Center. The upper records and inter-record times for the annual maximum earthquakes are:

$$\begin{array}{rcccccccccc} x_i & = & 6.0, & 6.5, & 7.2, & 7.4, & 7.6, & 7.9, & 8.0, & 8.3, & 8.4 \\ \tau_i & = & 3, & 3, & 1, & 15, & 10, & 28, & 39, & 26, & 1 \end{array}$$

We proceed as in Example 6.1. According to AIC, and the Log likelihood function, Gumbel distribution is more suitable than several other distributions (including Weibull, EVD, Frechet distributions) for modeling the previous data. The cdf of Gumbel distribution is of the form

$$F(y) = 1 - \exp\left[-e^{(y-\alpha)/\beta}\right], \quad -\infty < y < \infty,\ -\infty < \alpha < \infty,\ \beta > 0.$$

The data are analyzed in the following two cases:

1. In the first case, we suppose that the first 8 record-values have been observed. The prediction results in the first 3 rows of Table 6 are obtained via the MLE's $\hat{\alpha} = 6.59296$ and $\hat{\beta} = 1.00387$, which are computed from (6.1).
2. In the second case, all the first 9 record-values are assumed to be observed. An application to (6.1) again yields, $\hat{\alpha} = 6.58459$ and $\hat{\beta} = 1.07983$, which are very close to the MLE's computed from the complete data. The prediction results according to these estimates are shown in the last two rows of Table 6.

In such cases a point predictor based on P_1 is given by

$$\tilde{Y}^{(1)}(t_s) = Y^{(1)}(t_r) + E\left[Y^{(1)}(t_s)\right] - E\left[Y^{(1)}(t_r)\right], \; t_s > t_r,$$

where $E\left[Y^{(1)}(t_s)\right]$ and $E\left[Y^{(1)}(t_r)\right]$ are computed numerically.

Table 6: Point predictor and 95% upper limits for the next half record-value and the next record-value for the annual maximum earthquakes.

t_r	t_s	L	$\tilde{Y}^{(1)}(t_s)$	U_{P_1}	U_{P_2}
8	9	8.300	8.425	8.676	8.694
	9.5	8.300	8.537	8.776	8.799
	10	8.300	8.637	8.863	8.890
9	9.5	8.400	8.520	8.637	8.641
	10	8.400	8.628	8.759	8.767

Example 6.3 (One hour mean concentration of sulphur dioxide)**.** The following data represents the monthly maxima of 1 h mean concentration of sulphur dioxide in parts per hundred million (pphm) from Long Beach, California, during 1956 to 1974 for the month of October:

$$\begin{array}{cccccccccc} 26, & 14, & 27, & 15, & 16, & 16, & 11, & 10, & 14, & 12, \\ 15, & 40, & 29, & 13, & 20, & 41, & 31, & 28, & 11. & \end{array}$$

Roberts[44] shows that the Weibull model is a reasonably good for fitting this data. An application of extreme value $Q-Q$ plot by [16] supports Weibull model. The upper records and inter-record times for the above data are: $x_1 = 26,\; x_2 = 27,\; x_3 = 40,\; x_4 = 41$, and $\tau_1 = 2,\; \tau_2 = 9,\; \tau_3 = 4,\; \tau_4 = 1$. The MLE's of Weibull parameters based on the likelihood function (6.1) are $\hat{\alpha} = 2.3596$ and $\hat{\beta} = 24.5108$.

Based on the pivotal quantities P_1 and P_2, 90% PCI's for the next two records, respectively, are $(41, 52.328)$, $(41, 52.103)$ and $(41, 59.452)$, $(41, 59.382)$, which are shorter than the intervals obtained by [47] ((41.1590, 60.2449) and (41.9011, 75.5765)). Moreover, an unbiased point predictors for the next two record-values are obtained from (5.4), that is, $\tilde{Y}^{(1)}(5) = 45.344$ and $\tilde{Y}^{(1)}(6) = 49.187$.

7. CONCLUSION

In this article, we have proposed two predictive pivotal quantities for constructing prediction intervals of future ordinary (fractional) upper (lower) records from any continuous distribution. More details have been given for the exponential distribution. Prediction intervals constructed using this approach have been demonstrated, by using a simulation study and by applying it to real data. Example 6.3 shows that this method gives a shorter intervals than that given by Wang and Ye [47]. Moreover, the second case in the simulation study as well as the three real data examples show that, when the cdf of the data is unknown as always in practice, the given method is applicable with acceptable degree of accuracy. Also, it is noted that the coverage probability is closed to theoretical value $1 - \delta = 0.90$ and average upper (lower) limits of PCI are closed to expected values of upper (lower) limits based on both P_1 and P_2. Comparisons based on exact and estimated root mean square errors, indicate that the pivotal quantity P_1 is relatively better than P_2. Moreover, the root mean square errors, increase with increasing of the difference $t_s - t_r$. Finally, three real data sets have been completely analyzed.

A. ALGORITHMS

Based on the results of Rider [43], Rahman [41], Cramer [18] and Burkschat *et al.* [15], we can generate ordinary k-th upper (lower) record-values from any continuous cdf F with pdf f, by the following algorithm.

A.1. Algorithm 1

Step 1. Choose the values of n, k and determine the cdf F;

Step 2. Generate a random sample of size n from beta distribution, $Beta(k, 1)$, say $\mathcal{B}_1, \mathcal{B}_2, ..., \mathcal{B}_n$;

Step 3. Compute the k-th upper record-value $Y_r^{(k)}$, based on F by the formula

$$Y_r^{(k)} = F^{-1}\left(1 - \prod_{j=1}^{r} \mathcal{B}_j\right), \quad r = 1, 2, ..., n;$$

Step 4. Compute the k-th lower record-value $X_r^{(k)}$, based on F from the relation

$$X_r^{(k)} = F^{-1}\left(\prod_{j=1}^{r} \mathcal{B}_j\right), \quad r = 1, 2, ..., n.$$

The second algorithm relays on Theorem 1 and Definition 2 of Bieniek and Szynal [14]. The algorithm is formulated in a special case, whenever there is only a single fractional k-th upper record-value between two successive ordinary k-th upper record-values.

A.2. Algorithm 2

Step 1. Determine n, k and use Algorithm 1 to generate n ordinary k-th upper record-values, $W_i^{(k)}$, $i = 1, 2, ..., n$, based on EXP(1);

Step 2. Choose the real numbers $0 = t_0 < t_1 < ... < t_n$, such that, $i - 1 < t_i < i$, $\forall\ i = 1, ..., n$;

Step 3. Compute the fractional k-th upper record-values based on EXP(1) by Theorem 1 of Bieniek and Szynal (2004), that is,

$$W^{(k)}(t_i) = (1 - B_i^*)W_{[t_i]}^{(k)} + B_i^* W_{[t_i+1]}^{(k)}, \quad i = 1, 2, ..., n, \tag{A.1}$$

where $[t_i]$ denotes the greatest integer part of t_i, B_i^* is a random observation from beta distribution $Beta(t_i^*, 1 - t_i^*)$, independent of $W_{[t_i]}^{(k)}, i = 1, 2, ..., n$, and t_i^* denotes the fractional part of the numerical value of t_i;

Step 4. The fractional k-th upper record-values based on F, are then given by

$$Y^{(k)}(t_i) = F^{-1}\left(1 - e^{-W^{(k)}(t_i)}\right), \quad i = 1, 2, ..., n. \tag{A.2}$$

The general case can be accomplished by Theorems 2 and 3 of [14].

A.3. Algorithm 3

Step 1. Determine the number n of fractional upper records to be generated, the number of repetitions M, the real numbers $0 = t_0 < t_1 < ... < t_n$ with $i-1 < t_i < i,\ \forall\ i = 1, ..., n$ and the distribution with its parameter(s);

Step 2. Generate and store M arrays, each array include n of fractional k-th upper record-values;

Step 3. Determine the number of observed ordinary (fractional) k-th upper record-values r and the number of future ordinary (fractional) k-th upper record-value s, to be predicted;

Step 4. Find the numerical values of $p_i(\delta)$ by solving the nonlinear equations $F_{P_i}(p_i) = 1 - \delta,\ i = 1, 2$;

Step 5. Find the MLE's of the parameters based on the first r ordinary (fractional) k-th upper record-values;

Step 6. Compute the upper and lower limits for the PCI based on the pivotal quantities P_1 and P_2 by Theorems 3.1 and 3.2, and the point predictor(s) with

(**i**) the true values of parameters, and

(**ii**) the MLE's of parameters;

Step 7. Check whether, the observed value of $Y^{(k)}(t_s)$ did belong to the PCI;

Step 8. Repeat Steps 5, 6 and 7, M times;

Step 9. Compute the percentage of coverage probability, that is the percent that the true value of the future fractional record lies inside the PCI, the average of the lower and upper limits;

Step 10. Compute the root mean square errors and expected values of upper limits based on P_1 and P_2.

ACKNOWLEDGMENTS

The authors acknowledge the valuable suggestions from the referee and the associated editor for careful reading of the first version.

REFERENCES

[1] Ahsanullah, M. (1978). Record values and the exponential distribution, *Annals of the Institute of Statistical Mathematics*, **30**(1), 429–433.

[2] Ahsanullah, M. (1980). Linear prediction of record values for the two parameter exponential distribution, *Annals of the Institute of Statistical Mathematics*, **32**(1), 363–368.

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

[4] Al-Hussaini, E.K. and Ahmad, A.A. (2003). On Bayesian interval prediction of future records, *Test*, **12**(1), 79–99.

[5] Aly, A.E. (2015). Prediction and reconstruction of future and missing unobservable modified Weibull lifetime based on generalized order statistics, *Journal of the Egyptian Mathematical Society*, **24**(2), 309–318.

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

[7] Balakrishnan, N.; Ahsanullah, M. and Chan, P.S. (1995). On the logistic record values and associated inference, *Journal of Applied Statistical Science*, **2**(3), 233–248.

[8] Barakat, H.M.; El-Adll, M.E. and Amany, E.Aly. (2011). Exact prediction intervals for future exponential lifetime based on random generalized order statistics, *Computers & Mathematics with Applications*, **61**(5), 1366–1378.

[9] Barakat, H.M.; El-Adll, M.E. and Amany, E.Aly. (2014). Prediction intervals of future observations for a sample of random size from any continuous distribution, *Mathematics and Computers in Simulation*, **97**, 1–13.

[10] Barakat, H.M.; Nigm, E.M. and Aldallal, R.A. (2014a). Current records and record range with some applications, *Journal of the Korean Statistical Society*, **43**, 263–273.

[11] Barakat, H.M.; Nigm, E.M. and Aldallal, R.A. (2014b). Exact prediction intervals for future current records and record range from any continuous distribution, *SORT*, **38**(2), 251–270.

[12] Barakat, H.M.; Nigm, E.M.; El-Adll, M.E. and Yusuf, M. (2018). Prediction of future generalized order statistics based on exponential distribution with random sample size, *Statistical Papers*, **59**(2), 605–631.

[13] Berred, A.M. (1998). Prediction of record values, *Communications in Statistics – Theory and Methods*, **27**(9), 2221–2240.

[14] Bieniek, M. and Szynal, D. (2004). On the fractional record values, *Probability and Mathematical Statistics*, **24**(1), 27–46.

[15] Burkschat, M.; Cramer, E. and Kamps, U. (2003). Dual generalized order statistics, *Metron*, **61**(1), 13–26.

[16] Chan, P.S. (1993). *A statistical study of log-gamma distribution* (Ph.D. dissertation), McMaster University, Canada.

[17] Chandler, K.N. (1952). The distribution and frequency of record values, *Journal of the Royal Statistical Society: Series B*, **14**(2), 220–228.

[18] Cramer, E. (2003). *Contributions to Generalized Order Statistics*, Habililations-schrift, Reprint, University of Oldenburg.

[19] Doganaksoy, N. and Balakrishnan, N. (1997). A useful property of best linear unbiased predictors with applications to life-testing, *The American Statistician*, **51**(1), 22–28.

[20] Dunsmore, I.R. (1983). The future occurrence of records, *Annals of the Institute of Statistical Mathematics*, **35**(1), 267–277.

[21] Dziubdziela, W. and Kopociǹski, B. (1976). Limiting properties of the k-th record value, *Zastosowania Mathematyki*, **15**, 187–190.

[22] El-Adll, M.E. (2011). Predicting future lifetime based on random number of three parameters Weibull distribution, *Mathematics and Computers in Simulation*, **81**, 1842–1854.

[23] El-Adll, M.E. and Aly, A.E. (2016). Prediction intervals of future generalized order statistics from Pareto distribution, *Journal of Applied Statistical Science*, **22**(1–2), 111–125.

[24] Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics*, Second edition, Krieger, Malabar, Florida.

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

[26] Grudziefi, Z. and Szynal, D. (1983). *On the expected values of k-th record values and associated characterizations of distributions.* In "Probability and Statistical Decision Theory", Vol. A, 1985, (F. Konecny *et al.*, Eds.), Reidel, Dordrecht, pp. 119–127.

[27] Gulati, S. and Padgett, W.J. (2003). *Parametric and nonparametric inference from record-breaking data*, Springer-Verlag, New York.

[28] Hofmann, G. and Nagaraja, H.N. (2003). Fisher information in record data, *Metrika*, **57**, 177–193.

[29] Hoinkes, L.A. and Padgett, W. (1994). Maximum likelihood estimation from record-breaking data for the Weibull distribution, *Quality and Reliability Engineering International*, **10**(1), 5–13.

[30] Hutson, A.D. (1999). Calculating nonparametric confidence intervals for quantiles using fractional order statistics, *Journal Journal of Applied Statistic*, **26**(3), 343–353.

[31] Jones, M.C. (2002). On fractional uniform order statistics, *Statistics & Probability Letters*, **58**(1), 93–96.

[32] Kaminsky, K.S. and Nelson, P.I. (1975). Best Linear Unbiased Prediction of Order Statistics in Location and Scale Families, *Journal of the American Statistical Association*, **70**(349), 145–150.

[33] Kaminsky, K.S. and Nelson, P.I. (1998). *Prediction on order statistics.* In "Handbook of Statistics" (N. Balakrishnan and C.R. Rao, Eds.), 17, North Holland, Amsterdam, pp. 431–450.

[34] Kaminsky, K.S. and Rhodin, L.S. (1985). Maximum likelihood prediction, *Annals of the Institute of Statistical Mathematics*, **37**(1), 507–517.

[35] Kyzylaslan, F. and Nadar, M. (2015). Estimation with the generalized exponential distribution based on record values and inter-record times, *Journal of Statistical Computation and Simulation*, **85**(5), 978–999.

[36] Lawless, J.F. (1971). A prediction problem concerning samples from the exponential distribution with applications in life testing, *Technometrics*, **13**(4), 725–730.

[37] Nagaraja, H.N. (1984). Asymptotic linear prediction of extreme order statistics, *Annals of the Institute of Statistical Mathematics*, **36**, 289–299.

[38] Nagaraja, H.N. (1988). Record values and related statistics a review, *Communications in Statistics – Theory and Methods*, **17**(1), 2223–2238.

[39] Nagaraja, H.N. (1995). *Prediction problems.* In "The Exponential Distribution: Theory and Appl." (N. Balakrishnan and A.P. Basu, Eds.), pp. 139–163, Gordon and Breach, New York.

[40] Nevzorov, V.B. (2001). *Records: Mathematical Theory*, Translations of Mathematical Monographs, American Math. Soc., Providence, R.I.

[41] Rahman, N.A. (1964). Some generalisations of the distributions of product statistics arising from rectangular populations, *Journal of the American Statistical Association*, **59**(306), 557–563.

[42] Raqab, M. and Balakrishnan, N. (2008). Prediction intervals for future records, *Statistics & Probability Letters*, **78**(13), 1955–1963.

[43] Rider, P.R. (1955). The distribution of the product of maximum values in samples from a rectangular distribution, *Journal Journal of the American Statistical Association*, **50**(272), 1142–1143.

[44] Roberts, E. (1979). Review of statistics of extreme values with applications to air quality data: part II. Applications, *Journal Journal of the Air Pollution Control Association*, **29**(7), 733–740.

[45] Samaniego, F.J. and Whitaker, L.R. (1988). On estimating population characteristics from record-breaking observations. II: Nonparametric Results, *Naval Research Logistics Quarterly*, **35**(2), 221–236.

[46] Stigler, M.S. (1977). Fractional order statistics, with applications, *Journal Journal of the American Statistical Association*, **72**(359), 544–550.

[47] Wang, B.X. and Ye, Z. (2015). Inference on the Weibull distribution based on record values, *Computational Statistics & Data Analysis*, **83**, 26–36.