Statistics and Applications {ISSN 2454-7395 (online)}
Volume 21, No. 1, 2023 (New Series), pp 81–98

Nonparametric Prediction Intervals for Future Order Statistics and *k*-Record Values

Laji Muraleedharan and Manoj Chacko
Department of Statistics, University of Kerala, Trivandrum, India

Received: 03 February 2022; Revised: 25 April 2022; Accepted: 26 May 2022

Abstract

In this paper, we obtain distribution-free prediction intervals for future order statistics based on an observed sequence of k-record values. The Prediction intervals for future k-record values based on observed order statistics and prediction intervals of future record values based on observed k-record values are also derived in a similar manner. The coverage probabilities of the derived intervals are exact and independent of the parent distribution. Finally, two real data sets are used to illustrate the proposed methodologies developed in this paper.

Key words: Prediction intervals; Order statistics; Record values; k-Record values.

AMS Subject Classifications: Primary:62G30; Secondary: 62E15

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n arising from a population with absolutely continuous cumulative distribution function (cdf) $G(x)$ and probability density function (pdf) $g(x)$. By arranging the random sample in an increasing order of magnitude as $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$, the order statistics of the sample can be obtained. The ith order statistic of the sample $X_1, X_2, \ldots, X_n$ is then $X_{i:n}$. Order statistics have wide range of applications in many fields including industry, reliability analysis and material strength. For more discussions regarding the order statistics, one may refer to Arnold *et al.* (1992) and David and Nagaraja (2003). One major application of order statistics in the study of reliability of systems is the following. A system is called a *k-out-of-m* system if it consists of m components and the system functions satisfactorily if at least $k\,(\leq m)$ components function. If the lifetimes of the components are independently distributed, then the lifetime of the system coincides with that of the $(m-k+1)\,th$ order statistic of the lifetime of the components. Thus, order statistics play a key role in studying the lifetimes of such systems.

The cdf of the ith order statistic $X_{i:n}$ based on a random sample of size n from a continuous population with cdf $G(x)$ and pdf $g(x)$ is given by (see, Arnold *et al.*,1992)

$$F_{i:n}(x) = \sum_{r=i}^{n} \binom{n}{r} [G(x)]^r \left[\bar{G}(x)\right]^{n-r}, \ -\infty < x < \infty. \tag{1}$$

Corresponding Author: Laji Muraleedharan
Email: lajikmr@gmail.com

The pdf corresponding to the cdf (1) is given by

$$f_{i:n}(x) = \frac{1}{B(i, n-i+1)} \left[G(x)\right]^{i-1} \left[\bar{G}(x)\right]^{n-i} g(x), \; -\infty < x < \infty, \tag{2}$$

where $\bar{G} = 1 - G$ and $B\left(.,.\right)$ denotes the complete beta function.

Let $\{X_i, i \geq 1\}$ be a sequence of independent and identically distributed (iid) random variables with an absolutely continuous cdf $G(x)$ and pdf $g(x)$. If an observation X_j exceeds all of its previous observations, that is, $X_j > X_i$ for every $i < j$, then it is referred to as an upper record value. Thus X_1 is the first upper record value by definition. Similarly, the lower record values can be defined. Many authors have studied the record values of iid random variables as well as their features in the literature. Arnold *et al.* (1998), Ahsanullah (1995) and the literature referenced therein can be used to have a more in-depth look at this topic.

Since Chandler (1952) brought up the idea of record values for the first time in the literature, there has been a significant growth in the study of record values. Record values have many statistical applications, such as modelling and inference involving data pertaining to mining, sports, industry, seismology, life testing and so on. Interested Surveys are given in Glick (1978), Gulati and Padgett (1994), Ahsanullah (1995), Arnold *et al.* (1998), Nagaraja (1988) and the literature cited therein.

One of the challenges in dealing with problems involving inference with record data is that the expected waiting time for consecutive records after the first is infinite. Such an issue does not arise if we use the k-records proposed by Dziubdziela and Kopociński (1976). We use the following formal definition of k-record values given by Arnold *et al.* (1998).

For a fixed positive integer k, the upper k-record times $\tau_{n(k)}$ and the upper k-record values $U_{n(k)}$ are defined as follows.
Define $\tau_{1(k)} = k$ and $U_{1(k)} = X_{1:k}$ then for $n > 1$,

$$\tau_{n(k)} = min\left\{i : i > \tau_{n-1(k)}, \; X_i > X_{\tau_{n-1(k)}-k+1:\tau_{n-1(k)}}\right\}.$$

Then the sequence of upper k-record values $\left\{U_{n(k)}, n \geq 1\right\}$ is defined as

$$U_{n(k)} = X_{\tau_{n(k)}-k+1:\tau_{n(k)}}.$$

The cdf of the nth upper k-record value $U_{n(k)}$ for $n \geq 1$ is given by (see, Arnold *et al.*,1998)

$$F_{n(k)}(x) = 1 - \left[\bar{G}(x)\right]^k \sum_{i=1}^{n-1} \frac{\left[-k \log \bar{G}(x)\right]^i}{i!}, \quad -\infty < x < \infty. \tag{3}$$

The pdf corresponds to the cdf (3) is given by

$$f_{n(k)}(x) = \frac{k^n}{\Gamma(n)} \left[-\log \bar{G}(x)\right]^{n-1} \left[\bar{G}(x)\right]^{k-1} g(x), \; -\infty < x < \infty, \tag{4}$$

where $\Gamma(\cdot)$ denotes the complete gamma function. Similarly, we can define the lower k-record values as well.

For a fixed positive integer k, the sequence of lower k-record times $\left\{\tau^*_{n(k)}\right\}$ and lower k-record value $L_{n(k)}$ are defined as follows. Let
$\tau^*_{1(k)} = k$ and $L_{1(k)} = X_{k:k}$ then for $n > 1$,

$$\tau^*_{n(k)} = min\left\{j : j > \tau^*_{n-1(k)},\ X_j < X_{k:\tau^*_{n-1(k)}}\right\}.$$

Now the sequence of lower k-record values $\left\{L_{n(k)}, n \geq 1\right\}$ is defined by

$$L_{n(k)} = X_{k:\tau^*_{n(k)}}.$$

Recently, the k-records data has shown an increased trend in problems involving statistical inference and future event prediction. Chacko and Muraleedharan (2018) have obtained the Bayesian and maximum likelihood estimators for the parameters of a generalized exponential distribution based on k-record values. The same problem was discussed by Muraleedharan and Chacko (2019) for Gompertz distribution. The recurrence relation for the single and product moment of Gompertz distribution and its characterization based on k-records were studied by Minimol and Thomas (2014). The Bayesian estimation of parameters for a Gumbel distribution and the one sample prediction of future k-record values under the Bayesian frame work were studied by Malinowska and Szynal (2004). The best linear unbiased predictor (BLUP) for future k-record value based on k-records arising from a normal distribution was discussed by Chacko and Mary (2013) whereas the same problem for a generalized Pareto distribution was discussed by Muraleedharan and Chacko (2022). Paul and Thomas (2015) established some properties of upper k-record values which characterize the Weibull distribution and has derived the BLUP for the model. Deheuvels and Nevzorov (1994) studied the limiting behaviour of k-record values such as strong laws of large numbers, central limit theorems, functional laws of the iterated logarithm and strong invariance principles *etc.*

In statistical inference, predicting future events based on the current knowledge is a fundamental problem. It can be expressed in a variety of ways and in various settings. There are two different sorts of prediction problems. The one sample prediction problem is that the event to be predicted comes from the same sequence of events, whereas the two sample prediction problem is when the event to be predicted comes from a different independent sequence of events.

Several authors have considered prediction problem involving record values and order statistics. Hsieh (1997) developed the explicit expression for the prediction intervals for future Weibull order statistics. Al-Hussaini and Ahmad (2003) obtained the Bayesian prediction bounds for future record values from a general class of distributions. Prediction of distribution-free confidence intervals based on record values, order statistics and progressively type II censored samples were extensively discussed by Ahmadi and Balakrishnan (2005, 2008, 2010), Ahmadi *et al.* (2010) and Guilbaud (2004) respectively. In this paper, we consider the two sample distribution-free prediction intervals for order statistics and k-record values.

This paper is structured as follows. In Section 2, we use the observed k- record values to derive the prediction intervals and the corresponding prediction coefficient for future order statistics. In Section 3, based on the observed order statistics, we obtain the prediction intervals and its coefficient for the future k-record values. In Section 4, we consider the interval prediction of future record values based on observed k-record values. In Section 5, two real data sets are used to exemplify the proposed approaches presented in this paper and finally some concluding remarks are made in Section 6.

2. Prediction of order statistics based on k-record values

In this section, we consider the two-sided prediction intervals for an order statistic from the future sample based on the observed k-record values. Let $\left\{R_{i(k)}, i \geq 1\right\}$ be a sequence of observed upper (lower) k-record values arising from a population with absolutely continuous cdf $G(x)$. Suppose we are interested in obtaining an interval of the form $\left(R_{s(k)}, R_{t(k)}\right)$, $1 \leq s < t$, for the rth order statistic $Y_{r:n}$, $1 \leq r \leq n$, of the future sample of size n arising from the same population such that

$$P\left(R_{s(k)} \leq Y_{r:n} \leq R_{t(k)}\right) = 1 - \alpha.$$

Then the interval $\left(R_{s(k)}, R_{t(k)}\right)$ is called a $100\,(1-\alpha)\,\%$ prediction interval with prediction coefficient $(1-\alpha)$ for the future order statistic $Y_{r:n}$. In this section, we derive such two-sided prediction intervals for $Y_{r:n}$ with coverage probabilities that are free of the parent distribution function G.

2.1. Prediction of order statistics based on upper k-record values

Let $\{Y_i, i \geq 1\}$ be a sequence of iid random variables having an absolutely continuous cdf $G(x)$ and pdf $g(x)$. In the following theorem, we establish the prediction intervals for future order statistics based on the observed sequence of upper k-record values.

Theorem 1: Let $\left\{U_{i(k)}, i \geq 1\right\}$ be a sequence of observed upper k-record values arising from a population with absolutely continuous cdf G and pdf g. Let $Y_{1:n} \leq Y_{2:n} \leq \cdots \leq Y_{n:n}$ be the order statistics from a future random sample of size n arising from the same cdf G. Then $\left(U_{s(k)}, U_{t(k)}\right)$, for $1 \leq s < t$, is a prediction interval for the rth order statistic $Y_{r:n}$, for $1 \leq r \leq n$, whose coverage probability is free of G and is given by

$$\alpha_{1(k)}\left(s, t; r, n\right) = r\binom{n}{r}\sum_{i=s}^{t-1}\sum_{j=0}^{r-1}\binom{r-1}{j}\frac{(-1)^j\, k^i}{(n+k+j+1-r)^{i+1}}. \tag{5}$$

Proof: For a given real number v and for $1 \leq s < t$, we have

$$\begin{aligned} P\left(U_{s(k)} \leq v\right) &= P\left(U_{s(k)} \leq v, U_{t(k)} < v\right) + P\left(U_{s(k)} \leq v, U_{t(k)} \geq v\right) \\ &= P\left(U_{t(k)} < v\right) + P\left(U_{s(k)} \leq v \leq U_{t(k)}\right). \end{aligned}$$

Hence

$$P\left(U_{s(k)} \leq v \leq U_{t(k)}\right) = P\left(U_{s(k)} \leq v\right) - P\left(U_{t(k)} < v\right). \tag{6}$$

By using (3), (6) can be expressed as

$$P\left(U_{s(k)} \leq v \leq U_{t(k)}\right) = \left[\bar{G}(v)\right]^k \sum_{i=s}^{t-1} \frac{\left[-k \log \bar{G}(v)\right]^i}{i!}. \tag{7}$$

Now for $1 \leq s < t$, and using the conditioning arguments, we can write (7) as

$$\begin{aligned}
\alpha_{1(k)}(s,t;r,n) &= P\left(U_{s(k)} \leq Y_{r:n} \leq U_{t(k)}\right) \\
&= \int_{-\infty}^{\infty} P\left(U_{s(k)} \leq Y_{r:n} \leq U_{t(k)} | Y_{r:n} = v\right) f_{r:n}(v) dv \\
&= \int_{-\infty}^{\infty} P\left(U_{s(k)} \leq v \leq U_{t(k)}\right) f_{r:n}(v) dv \\
&= \sum_{i=s}^{t-1} \frac{n!}{i!\,(r-1)!\,(n-r)!} \int_{-\infty}^{\infty} \left[-k \log \bar{G}(v)\right]^i \left[\bar{G}(v)\right]^{n+k-r} \left[G(v)\right]^{r-1} \\
&\times \; g(v) dv.
\end{aligned} \tag{8}$$

Taking $y = -k \log \bar{G}(v)$ and applying the binomial expansion, (8) reduces to the following

$$\begin{aligned}
\alpha_{1(k)}(s,t;r,n) &= \frac{r}{k}\binom{n}{r} \sum_{i=s}^{t-1}\sum_{j=0}^{r-1} \frac{(-1)^j}{i!}\binom{r-1}{j} \int_{y=0}^{\infty} y^i \exp\left[-\left(\frac{n+k+j+1-r}{k}\right) y\right] dy \\
&= r\binom{n}{r} \sum_{i=s}^{t-1}\sum_{j=0}^{r-1} \binom{r-1}{j} \frac{(-1)^j k^i}{(n+k+j+1-r)^{i+1}}.
\end{aligned} \tag{9}$$

Hence the proof. □

If n, r and the desired confidence level α_0 are supplied, we can choose s and t so that $\alpha_{1(k)}(s,t;r,n)$ surpasses α_0 . Since $\alpha_{1(k)}(s,t;r,n)$ is a step function, the confidence coefficient may not equal to α_0 but may be set to a value somewhat higher than α_0. Furthermore, the choice of s and t is not unique. So, for a given confidence level α_0, r and n, we would like to construct a prediction interval whose expected length as short as possible among all prediction intervals with the same level. First, notice that the two-sided prediction intervals exist for a given α_0, r and n if and only if, for large m,

$$P\left(U_{1(k)} \leq Y_{r:n} \leq U_{m(k)}\right) \geq \alpha_0.$$

We have evaluated $\alpha_{1(k)}(s,t;r,n)$ for $n = 20, 30$ and some selected values of (s,t) and r for $k = 2$ and $k = 3$ and the values are presented in Table 1. It can be observed that the prediction coefficient is increasing in r when the other parameters (s,t) and n are fixed and achieves reasonable prediction coefficient value when r close to n. It is also observed that for fixed n, r and k, the prediction coefficient $\alpha_{1(k)}(s,t;r,n)$ is decreasing in s and increasing in t.

2.2. Prediction of order statistics based on lower k-record values

In this subsection, we consider the prediction intervals for future order statistics on the basis of the observed lower k-record values. If $L_{n(k)}$ denotes the nth lower k-record value, then the cdf of $L_{n(k)}$ is given by

$$F^*_{n(k)}(x) = [G(x)]^k \sum_{s=1}^{n-1} \frac{[-k \log G(x)]^s}{s!}, \quad -\infty < x < \infty. \tag{10}$$

The pdf corresponds to the cdf (10) is given by

$$f^*_{n(k)}(x) = \frac{k^n}{\Gamma(n)} [-\log G(x)]^{n-1} [G(x)]^{k-1} g(x), \ -\infty < x < \infty. \tag{11}$$

Now we can establish the following theorem for the interval prediction of future order statistics based on the observed sequence of lower k-record values.

Theorem 2: Suppose the conditions of Theorem 1 hold and let $\left\{L_{i(k)}, i \geq 1\right\}$ be the sequence of observed lower k-record values emerging from the population. Then $\left(L_{t(k)}, L_{s(k)}\right)$, for $1 \leq s < t$, is a prediction interval for the rth order statistic $Y_{r:n}$, for $1 \leq r \leq n$, whose prediction coefficient is free of G and is given by

$$\alpha_{2(k)}(s,t;r,n) = r\binom{n}{r} \sum_{i=s}^{t-1} \sum_{j=0}^{n-r} \binom{n-r}{j} \frac{(-1)^j k^i}{(r+k+j)^{i+1}}. \tag{12}$$

Proof: The proof is similar to that of Theorem 1 and thus omitted. □

Remark 1: Since $\alpha_{2(k)}(s,t;r,n) = \alpha_{1(k)}(s,t;n-r+1,n)$, we can use Table 1 for evaluating (12).

2.3. Prediction of order statistics based on upper and lower k-record values jointly

In certain studies such as meteorological studies, the upper and lower k-record values are observed simultaneously. In such studies, when predicting the order statistics from a future sample, it is preferable to examine both the upper and lower k-record values together.

Theorem 3: Suppose the conditions of Theorem 1 hold; let $L_{s(k)}$ and $U_{t(k)}$ denote the sth lower k-record and tth upper k-record values respectively. Then $\left(Ł_{s(k)}, U_{t(k)}\right)$ is a prediction interval for the rth order statistic $Y_{r:n}$, for $1 \leq r \leq n$, with coverage probability free of G and is given by

$$\alpha_{3(k)}(s,t;r,n) = r\binom{n}{r}\left[\sum_{i=0}^{s-1}\sum_{j=0}^{n-r} \frac{(-1)^j \binom{n-r}{j} k^i}{(j+k+r)^{i+1}} + \sum_{i=0}^{t-1}\sum_{j=0}^{r-1} \frac{(-1)^j \binom{r-1}{j} k^i}{(n+j+k+1-r)^{i+1}}\right] - 1. \tag{13}$$

Proof: For a fixed real number v and $1 \leq s < t$, we can express

$$P\left(L_{s(k)} \leq v \leq U_{t(k)}\right) = [G(v)]^k \sum_{i=0}^{s-1} \frac{[-k \log G(v)]^i}{i!} + \left[\bar{G}(v)\right]^k \sum_{i=0}^{t-1} \frac{\left[-k \log \bar{G}(v)\right]^i}{i!} - 1. \tag{14}$$

Now for $1 \leq s < t$, and using the conditioning arguments, we can write (14) as

$$
\begin{aligned}
\alpha_{3(k)}(s,t;r,n) &= P\left(L_{s(k)} \leq Y_{r:n} \leq U_{t(k)}\right) \\
&= \int_{-\infty}^{\infty} P\left(U_{s(k)} \leq Y_{r:n} \leq U_{t(k)} | Y_{r:n} = v\right) f_{r:n}(v)dv \\
&= \int_{-\infty}^{\infty} P\left(U_{s(k)} \leq v \leq U_{t(k)}\right) f_{r:n}(v)dv \\
&= \sum_{i=1}^{s-1} \frac{n!}{i!\,(n-r)!\,(r-1)!} \int_{-\infty}^{\infty} [-k\log G(v)]^{i} \left[\bar{G}(v)\right]^{n-r} [G(v)]^{k+r-1} g(v)dv \\
&+ \sum_{i=1}^{t-1} \frac{n!}{i!\,(n-r)!\,(r-1)!} \int_{-\infty}^{\infty} \left[-k\log \bar{G}(v)\right]^{i} \left[\bar{G}(v)\right]^{k+n-r} [G(v)]^{r-1} g(v)dv - 1 \\
&= r\binom{n}{r} \sum_{i=1}^{s-1}\sum_{j=0}^{n-r} \frac{(-1)^{j}\binom{n-r}{j}}{i!k} \int_{y=0}^{\infty} y^{i}\left(e^{-\frac{y}{k}}\right)^{j+k+r} dy + r\binom{n}{r} \sum_{i=1}^{t-1}\sum_{j=0}^{r-1} \frac{(-1)^{j}\binom{r-1}{j}}{i!k} \\
&\times \int_{z=0}^{\infty} z^{i}\left(e^{-\frac{z}{k}}\right)^{n+j+k+1-r} dz - 1 \\
&= r\binom{n}{r}\left[\sum_{i=0}^{t-1}\sum_{j=0}^{r-1} \frac{(-1)^{j}\binom{n-r}{j}k^{i}}{(j+k+r)^{i+1}} + \sum_{i=0}^{t-1}\sum_{j=0}^{r-1} \frac{(-1)^{j}\binom{r-1}{j}k^{i}}{(n+j+k+1-r)^{i+1}}\right] - 1.
\end{aligned}
$$

Hence the proof. □

Table 2 provides the values of $\alpha_{3(k)}(s,t;r,n)$ for $n = 10, 20$ and 30 and some selected values of (s,t) and r for $k = 2$ and $k = 3$. We can see that the prediction coefficient improves when the intervals are constructed upper and lower k-record values jointly. It is also observed that for fixed n, r and k, the prediction coefficient $\alpha_{3(k)}(s,t;r,n)$ is non-decreasing in s and t.

3. Prediction of future k-record values based on order statistics

Suppose we are interested in obtaining an interval for the rth future k-record value $R_{r(k)}$ (upper or lower) based on the observed order statistics of size n of the form $(X_{s:n}, X_{t:n})$, $1 \leq s < t \leq n$, such that

$$P\left(X_{s:n} \leq R_{r(k)} \leq X_{t:n}\right) = 1 - \alpha.$$

Then we refer the interval $(X_{s:n}, X_{t:n})$ as a $100(1-\alpha)\,\%$ prediction interval with prediction coefficient $(1-\alpha)$ for the rth future k- record value $R_{r(k)}$. In this section, we derive such two-sided prediction intervals with coverage probabilities being free of the parent distribution.

3.1. Prediction of upper *k*- record values based on order statistics

In this subsection, we wish to predict the rth future upper k-record value $U_{r(k)}$ based on the observed order statistics.

Theorem 4: Let $Y_{1:n} \leq Y_{2:n} \leq \cdots \leq Y_{n:n}$ be the observed order statstics arising from a random sample of size n from a population with absolutely continuous cdf G and pdf g respectively. Then $(Y_{s:n}, Y_{t:n})$, for $1 \leq s < t \leq n$, is a prediction interval for the rth future upper k-record value $U_{r(k)}$ arising from the same population whose coverage probability is free of G and is given by

$$\alpha_{4(k)}\left(s,t;r,n\right) = \sum_{i=s}^{t-1}\sum_{j=0}^{i}\binom{n}{i}\binom{i}{j}\frac{\left(-1\right)^{j}k^{r}}{\left(n+j+k-i\right)^{r}}. \tag{15}$$

Proof: For any real number v and $1 \leq s < t \leq n$, by using (1), we obtain the following

$$P\left(Y_{s:n} \leq v \leq Y_{t:n}\right) = \sum_{i=s}^{t-1}\binom{n}{i}\left[G(v)\right]^{i}\left[\bar{G}(v)\right]^{n-i}. \tag{16}$$

Now for $1 \leq s < t \leq n$, and using the conditioning arguments, we can write

$$\begin{aligned}
\alpha_{4(k)}\left(s,t;r,n\right) &= P\left(Y_{s:n} \leq U_{r(k)} \leq Y_{t:n}\right)\\
&= \int_{-\infty}^{\infty} P\left(X_{s:n} \leq v \leq X_{t:n} | U_{r(k)} = v\right) f_{r(k)}(v)dv\\
&= \int_{-\infty}^{\infty} P\left(X_{s:n} \leq v \leq X_{t:n}\right) f_{r(k)}(v)dv\\
&= \sum_{i=s}^{t-1}\binom{n}{i}\frac{k^{r}}{(r-1)!}\int_{-\infty}^{\infty}\left[-\log\bar{G}(v)\right]^{r-1}\left[\bar{G}(v)\right]^{n+k-i-1}\left[G(v)\right]^{i}g(v)dv\\
&= \sum_{i=s}^{t-1}\sum_{j=0}^{i}\binom{n}{i}\binom{i}{j}\frac{k^{r}\left(-1\right)^{j}}{(r-1)!}\int_{-\infty}^{\infty} y^{r-1}\exp\left[-\left(n+j+k-i\right)y\right]dy\\
&= \sum_{i=s}^{t-1}\sum_{j=0}^{i}\binom{n}{i}\binom{i}{j}\frac{\left(-1\right)^{j}k^{r}}{\left(n+j+k-i\right)^{r}}.
\end{aligned}$$

Hence the proof. □

For a given confidence level α_0 and specified r, we would like to construct prediction interval whose expected length as short as possible among all prediction intervals with the same confidence level. First observe that, for a given α_0 and r, the two-sided prediction interval exists if and only if

$$P\left(X_{1:n} \leq U_{r(k)} \leq X_{n:n}\right) \geq \alpha_0.$$

Table 3 represents the values of $\alpha_{4(k)}\left(s,n;r,n\right)$ for $n = 10, 20, 30, 35$ and 40 and some selected values of s and r for $k = 2$ and $k = 3$. We can observe that $\alpha_{4(k)}\left(s,n;r,n\right)$ is decreasing in r and s but improves with n and k.

3.2. Prediction of lower k-record values based on order statistics

For predicting lower k-record values, we consider an interval $(X_{s:n}, X_{t:n})$, for $1 \leq s < t \leq n$, based on the observed order statistics. Analogous to the result presented for upper k-record values, we obtain the following theorem.

Theorem 5: Suppose the conditions of Theorem 4 hold; then $(Y_{s:n}, Y_{t:n})$, for $1 \leq s < t \leq n$, is a prediction interval for the rth future lower k-record value $L_{r(k)}$ arising from the same population whose coverage probability is free of G and is given by

$$\alpha_{5(k)}\left(s,t;r,n\right) = \sum_{i=s}^{t-1}\sum_{j=0}^{n-i}\binom{n}{i}\binom{n-i}{j}\frac{\left(-1\right)^{j}k^{r}}{\left(i+j+k\right)^{r}}. \tag{17}$$

Proof: Proof is similar to that of Theorem 4 and hence omitted. □

Remark 2: Note that if

$$\alpha_{4(k)}\left(s,t;r,n\right) = \sum_{i=s}^{t-1}\binom{n}{i}\psi_{k}\left(i,j,r;n\right)$$

then

$$\alpha_{5(k)}\left(s,t;r,n\right) = \sum_{i=s}^{t-1}\binom{n}{i}\psi_{k}\left(n-i,j,r;n\right),$$

where

$$\psi_{k}\left(i,j,r;n\right) = \sum_{j=0}^{i}\binom{i}{j}\frac{\left(-1\right)^{j}k^{r}}{\left(n+j+k-i\right)^{r}}. \tag{18}$$

Thus we can use Table 3 for evaluating (17) by making a simple modification.

4. Prediction of future record values based on k-record values

Let $\left\{R_{i(k)}, i \geq 1\right\}$ be a sequence of observed k-record values arising from a population with absolutely continuous cdf $G(x)$. Suppose we are interested in obtaining an interval for the rth future record value R_r of the form $\left(R_{m(k)}, R_{n(k)}\right), 1 \leq m < n$, such that

$$P\left(R_{n(k)} \leq R_r \leq R_{n(k)}\right) = 1-\alpha.$$

Then we refer the interval $\left(R_{m(k)}, R_{n(k)}\right)$ as a $100\left(1-\alpha\right)\%$ prediction interval with prediction coefficient $(1-\alpha)$ for the future record value R_r. In this section, we derive such two-sided prediction intervals for R_r with coverage probabilities being free of the parent distribution G.

4.1. Prediction of future upper record values based on upper k-record values

In this subsection, we wish to predict the rth future upper record value based on the observed sequence of upper k-record values.

Theorem 6: Let $\left\{U_{i(k)}, i \geq 1\right\}$ be a sequence of observed upper k- record values arising from a population with absolutely continuous cdf G. Then $\left(U_{s(k)}, U_{t(k)}\right)$, for $1 \leq s < t$, is a prediction interval for the rth future upper record value U_r arising from the same population with the corresponding prediction coefficient is given by

$$\alpha_{6(k)}\left(s,t;r\right) = \sum_{j=s}^{t-1} \binom{j+r-1}{j} \frac{k^j}{(1+k)^{j+r}}. \tag{19}$$

Proof: For a given real number v and for $1 \leq s < t$, we can express

$$P\left(U_{s(k)} \leq v \leq U_{t(k)}\right) = \left[\bar{G}(v)\right]^k \sum_{j=s}^{t-1} \frac{\left[-k \log \bar{G}(v)\right]^j}{j!}. \tag{20}$$

Now for $s < t$, and using the conditioning arguments, we can write (20) as

$$\begin{aligned}
\alpha_{6(k)}\left(s,t;r\right) &= P\left(U_{s(k)} \leq U_r \leq U_{t(k)}\right) \\
&= \int_{-\infty}^{\infty} P\left(U_{s(k)} \leq U_r \leq U_{t(k)} | U_r = v\right) f_{r(1)}(v) dv \\
&= \int_{-\infty}^{\infty} P\left(U_{s(k)} \leq v \leq U_{t(k)}\right) f_{r(1)}(v) dv \\
&= \sum_{j=s}^{t-1} \frac{1}{j!\,(r-1)!} \int_{-\infty}^{\infty} \left[-k \log \bar{G}(v)\right]^j \left[-\log \bar{G}(v)\right]^{r-1} \left[\bar{G}(v)\right]^k g(v) dv \\
&= \sum_{j=s}^{t-1} \binom{j+r-1}{j} \frac{k^j}{(1+k)^{j+r}}.
\end{aligned}$$

Hence the proof. □

Let W denote a negative binomial random variable counting the number of trials needed to get rth success where the success probability $p = 1/\left(k+1\right)$. Then, the expression in (19) can be viewed as the probability that the rth success occurs between sth and $t-1$ trials; that is, it represents $P\left(s \leq W < t\right)$, and hence $\alpha_{6(k)}\left(s,t;r\right)$ can be directly computed from negative binomial cdf using common statistical packages.

4.2. Prediction of future lower record values based on lower *k*-record values

In this section, we construct the prediction intervals for future lower record value based on the observed sequence lower k- record values. Analogous to the result presented for upper record values, we obtain the following theorem.

Theorem 7: Suppose the conditions of Theorem 6 hold, let $\left\{L_{n(k)}, n \geq 1\right\}$ be a sequence of observed lower k-record values arising from a population. Then $\left(L_{t(k)}, L_{s(k)}\right)$, for $1 \leq s < t$, is a prediction interval for the rth future lower record value L_r arising from the same population with the corresponding prediction coefficient is given by (19).

Proof: Proof is similar to that of Theorem 6 and hence omitted. □

4.3. Prediction of upper record value based on lower and upper k-record values jointly

There are some situations wherein upper and lower k-record values are observed jointly, just as in the case of weather data. In these cases, it would be better to consider the upper and lower k-record values jointly to predict the future upper record value of a future sample.

Theorem 8: Let $\left\{L_{i(k)}, i \geq 1\right\}$ and $\left\{U_{i(k)}, i \geq 1\right\}$ respectively denote the observed sequences of lower and upper k-record values arising from a population with absolutely continuous cdf G. Then $\left(L_{s(k)}, U_{t(k)}\right)$, for $1 \leq s < t$, is a prediction interval for the rth future upper record value U_r of the future random sample arising from the same population with the corresponding prediction coefficient, denoted by $\alpha_{7(k)}\left(s, t; r, n\right)$ being free of G; it can be expressed as

$$\alpha_{7(k)}\left(s, t; r\right) = \sum_{j=1}^{s-1} \frac{\theta_k\left(j, r\right)}{j!\left(r-1\right)!} + \alpha_{6(k)}\left(0, t; r\right) - 1, \tag{21}$$

where

$$\theta_k\left(j, r\right) = \int_0^1 y^k \left(-k\log y\right)^j \left[-\log\left(1-y\right)\right]^{r-1} dy \tag{22}$$

and $\alpha_{6(k)}\left(0, t, r\right)$ is defined by (19).

Proof: For a given real number v and for $1 \leq s < t$, we obtain

$$\begin{aligned} P\left(L_{s(k)} \leq v \leq U_{t(k)}\right) &= P\left(L_{s(k)} \leq v\right) - P\left(U_{t(k)} \leq v\right) \\ &= \sum_{j=0}^{s-1} \frac{\left[-k\log G(v)\right]^j}{j!}\left[G(v)\right]^k + \sum_{j=0}^{t-1} \frac{\left[-k\log \bar{G}(v)\right]^j}{j!}\left[\bar{G}(v)\right]^k - 1. \end{aligned} \tag{23}$$

Now for $1 \leq s < t$, and using the conditioning arguments, we can write (23) as

$$\begin{aligned} \alpha_{7(k)}\left(s, t; r\right) &= P\left(L_{s(k)} \leq U_r \leq U_{t(k)}\right) \\ &= \int_{-\infty}^{\infty} P\left(L_{s(k)} \leq U_r \leq U_{t(k)} | U_r = v\right) g_r(v) dv \\ &= \int_{-\infty}^{\infty} P\left(L_{s(k)} \leq v \leq U_{t(k)}\right) g_r(v) dv \\ &= \sum_{j=0}^{s-1} \frac{1}{j!\left(r-1\right)!} \int_{-\infty}^{\infty} \left[-k\log G(v)\right]^j \left[-\log \bar{G}(v)\right]^{r-1} \left[G(v)\right]^k g(v) dv \\ &+ \sum_{j=0}^{t-1} \frac{1}{j!\left(r-1\right)!} \int_{-\infty}^{\infty} \left[-k\log \bar{G}(v)\right]^j \left[-\log \bar{G}\right]^{r-1} \left[\bar{G}(v)\right]^k g(v) dv - 1. \end{aligned} \tag{24}$$

Taking $y = G(v)$ in the first integral and $z = -k \log \bar{G}(v)$ in the second integral of (24) and then evaluating, we obtain

$$\alpha_{7(k)}(s,t;r) = \sum_{j=0}^{s-1} \frac{\theta_k(j,r)}{j!\,(r-1)!} + \sum_{j=0}^{t-1} \frac{k^j}{(1+k)^{j+r}} \binom{j+r-1}{j} - 1, \tag{25}$$

where $\theta_k(j,r)$ is defined in (22). Therefore the prediction interval for the rth upper record value U_r from the future sequence is $\left(L_{s(k)}, U_{t(k)}\right)$ whose prediction coefficient is free of the parent distribution G and is given by

$$\alpha_{7(k)}(s,t;r) = \sum_{j=1}^{s-1} \frac{\theta_k(j,r)}{j!\,(r-1)!} + \alpha_{6(k)}(0,t;r) - 1. \tag{26}$$

Hence the proof. □

Tables 4 and 5 represent the values of $\alpha_{7(k)}(s,t;r)$ when $r = 1$ and $r = 2$ for $k = 2$ and $k = 3$ with $1 \le s \le 7$ and $4 \le t \le 7$.

5. Illustration using real data

Example 1: We use the data set given in Arnold *et al.*(1998, pp.49-50) which represent the average July temparatures (in degrees centigrade) of Neuenberg, Switzerland, during the period 1864-1993, and extract the *2*- record values to illustrate the prediction methods described for predicting future order statistics. Ahmadi and Balakrishnan (2011) used the same data set for predicting future order statistics based on observed ordinary record values. The first order autocorrelation for the data set at the first three lags are 0.022, -0.007 and -0.076 respectively. This small amount of autocorrelation shows that the data is independent in nature. The upper and lower *2*- record values extracted from the data set are obtained as given below.

m	1	2	3	4	5	6	7	8	9	10
$U_{m(2)}$	19.0	19.7	20.1	21.0	21.4	21.7	22	22.1	22.3	22
$L_{m(2)}$	20.1	19	18.4	17.4	17.2	16.2	15.8	15.6	-	-

Based on the observed upper and lower 2-record values and by using Table 2, we obtain the prediction intervals of future order statistics with prediction coefficient at least 90% for $n = 10, 20$ and 30 are presented in the following table.

(n,r)	(s,t)	(L_s,U_t)	$\alpha_{3(2)}(s,t;r,n)$	(n,r)	(s,t)	(L_s,U_t)	$\alpha_{3(2)}(s,t;r,n)$
$(10,6)$	$(5,7)$	$(17.2,22)$	0.9705	$(20,15)$	$(4,5)$	$(17.4,21.4)$	0.9456
$(10,8)$	$(3,8)$	$(18.4,22.1)$	0.9279	$(30,10)$	$(7,4)$	$(15.8,21)$	0.9726
$(20,5)$	$(7,4)$	$(15.8,21)$	0.9433	$(30,20)$	$(8,7)$	$(15.6,22)$	0.9900

When comparing the results in Table 2 to those of Ahmadi and Balakrishnan (2010), we see that when upper and lower record values are evaluated jointly, the interval prediction coefficient increases with lower values of k.

Example 2: Consider the amount of annual rainfall at Los Angeles Civic Centre (LACC) during 1900-2000. Then by Ahmadi and Balakrishnan (2011), the order statistics corresponding to the data set is given by

r	1	2	3	4	5	6	7	8	9	10
$Year$	1960	1958	1923	1971	1975	1947	1989	1986	1969	1963
$Y_{r:n}$	4.85	5.58	6.67	7.17	7.21	7.22	7.35	7.66	7.74	7.93
r	20	30	50	70	80	90	95	98	99	100
$Year$	1980	1941	1928	1926	1921	1937	1992	1982	1940	1977
$Y_{r:n}$	8.96	11.10	12.66	18.03	19.66	23.43	27.36	31.28	32.76	33.44

By using Table 3, we obtain the prediction intervals of future k-record values with prediction coefficient at least 90% for $k = 2$ and $k = 3$ are presented in the following table.

(n, r)	s	$(Y_{s:n}, Y_{n:n})$	$\alpha_{4(2)}(s, n; r, n)$	(n, r)	s	$(Y_{s:n}, Y_{n:n})$	$\alpha_{4(3)}(s, n; r, n)$
$(35, 4)$	6	$(7.21, 33.44)$	0.9204	$(20, 4)$	6	$(7.21, 33.44)$	0.9331
$(35, 4)$	8	$(7.66, 33.44)$	0.9185	$(40, 5)$	10	$(7.93, 33.44)$	0.9733
$(40, 4)$	10	$(7.93, 33.44)$	0.9288	$(40, 6)$	20	$(8.96, 33.44)$	0.9273

Ahmadi and Balakrishnan (2010) also used the same data set for constructing prediction intervals for future ordinary record values.

6. Conclusion

In this paper, we derived the distribution-free prediction intervals for order statistics and record values based on observed k-record values, as well as for future k-record values based on observed order statistics. The coverage probabilities of these intervals are exact and independent of the parent distribution. The proposed method can be extended to develop outer and inner prediction intervals for future k-record intervals.

Acknowledgement

The authors would like to thank the Chair Editor and the anonymous reviewer for their valuable comments and suggestions that significantly improved the quality of the paper.

Table 1: The values of $\alpha_{1(k)}(s,t;r,n)$ for $n=20$ and 30 and some selected values of s,t and r when $k=2$ and $k=3$

n	r	s	$k=2$				$k=3$			
			t				t			
			4	6	8	10	4	6	8	10
20	5	1	0.4061	0.4111	0.4113	0.4113	0.5208	0.5381	0.5392	0.5392
		2	0.1103	0.1152	0.1154	0.1154	0.1902	0.2075	0.2086	0.2086
		3	0.0210	0.0260	0.0261	0.0261	0.0478	0.0651	0.0661	0.0662
	10	1	0.6562	0.7089	0.7139	0.7143	0.6789	0.8104	0.8349	0.8381
		2	0.3204	0.3732	0.3782	0.3785	0.4104	0.5419	0.5664	0.5697
		3	0.1028	0.1556	0.1606	0.1609	0.1644	0.2958	0.3203	0.3236
	12	1	0.6895	0.7894	0.8036	0.8051	0.6274	0.8358	0.8936	0.9049
		2	0.3927	0.4926	0.5068	0.5083	0.4277	0.6362	0.6939	0.7052
		3	0.1468	0.2467	0.2609	0.2624	0.1951	0.4036	0.4613	0.4726
	15	1	0.6327	0.8414	0.8971	0.9074	0.4378	0.7435	0.8969	0.9501
		2	0.4331	0.6418	0.6975	0.7078	0.3413	0.6471	0.8004	0.8537
		3	0.1978	0.4065	0.4622	0.4725	0.1837	0.4894	0.6428	0.6960
	17	1	0.5013	0.7894	0.9112	0.9468	0.2635	0.5639	0.7937	0.9152
		2	0.3794	0.6676	0.7893	0.8249	0.2199	0.5204	0.7502	0.8717
		3	0.1961	0.4842	0.6059	0.6415	0.1305	0.4309	0.6607	0.7822
	19	1	0.2795	0.5792	0.8001	0.9151	0.0959	0.2817	0.5125	0.7129
		2	0.2313	0.5309	0.7518	0.8668	0.0847	0.2705	0.5014	0.7017
		3	0.1354	0.4351	0.6560	0.7710	0.0552	0.2410	0.4718	0.6722
	20	1	0.1384	0.3555	0.5859	0.7633	0.0342	0.1238	0.2729	0.4487
		2	0.1195	0.3365	0.5669	0.7443	0.0309	0.1206	0.2697	0.4455
		3	0.0749	0.2919	0.5223	0.6998	0.0211	0.1108	0.2599	0.4357
30	20	1	0.6908	0.8379	0.8637	0.8667	0.5546	0.8264	0.9214	0.9433
		2	0.4366	0.5836	0.6095	0.6124	0.4117	0.6835	0.7785	0.8004
		3	0.1806	0.3277	0.3535	0.3565	0.2065	0.4783	0.5733	0.5952
	22	1	0.6498	0.8527	0.9008	0.9083	0.4523	0.7661	0.9125	0.9574
		2	0.4449	0.6477	0.6959	0.7034	0.3553	0.6691	0.8155	0.8604
		3	0.2017	0.4045	0.4527	0.4602	0.1920	0.5059	0.6523	0.6972
	24	1	0.5727	0.8359	0.9219	0.9402	0.3323	0.6591	0.8652	0.9505
		2	0.4213	0.6846	0.7705	0.7889	0.2741	0.6009	0.8070	0.8923
		3	0.2086	0.4718	0.5578	0.5762	0.1590	0.4858	0.6918	0.7771
	26	1	0.4538	0.7676	0.9131	0.9575	0.2069	0.4992	0.7545	0.9013
		2	0.3565	0.6703	0.8158	0.8602	0.1781	0.4704	0.7257	0.8725
		3	0.1926	0.5064	0.6519	0.6963	0.1105	0.4029	0.6581	0.8049
	28	1	0.2900	0.6077	0.8312	0.9363	0.0936	0.2907	0.5402	0.7506
		2	0.2423	0.5599	0.7834	0.8885	0.0837	0.2807	0.5303	0.7407
		3	0.1433	0.4609	0.6844	0.7895	0.0556	0.2526	0.5022	0.7126
	29	1	0.1930	0.4699	0.7216	0.8770	0.0489	0.1777	0.3804	0.5949
		2	0.1661	0.4429	0.6947	0.8501	0.0445	0.1733	0.3760	0.5905
		3	0.1031	0.3800	0.6317	0.7872	0.0306	0.1594	0.3621	0.5766
	30	1	0.0911	0.2718	0.4982	0.6964	0.0166	0.0730	0.1877	0.3465
		2	0.0808	0.2615	0.4879	0.6861	0.0154	0.0718	0.1864	0.3452
		3	0.0529	0.2336	0.4601	0.6582	0.0109	0.0674	0.1820	0.3408

Table 2: The values of $\alpha_{3(k)}(s,t;r,n)$ for $n=10,20$ and 30 and some selected values of s,t and r when $k=2$ and $k=3$

n	r	s	$k=2$					$k=3$				
			t					t				
			4	5	6	7	8	4	5	6	7	8
10	2	2	0.1569	0.1601	0.1607	0.1609	0.1609	0.0060	0.0156	0.0184	0.0192	0.0194
		3	0.3220	0.3252	0.3259	0.3260	0.3260	0.0296	0.0392	0.0420	0.0428	0.0430
		5	0.6527	0.6559	0.6565	0.6567	0.6567	0.0861	0.0957	0.0985	0.0993	0.0995
	4	2	0.3701	0.3909	0.3974	0.3992	0.3998	0.0192	0.0693	0.0905	0.0990	0.1022
		3	0.6046	0.6253	0.6318	0.6337	0.6342	0.0976	0.1477	0.1689	0.1774	0.1806
		5	0.8715	0.8922	0.8987	0.9006	0.9011	0.2125	0.2625	0.2838	0.2922	0.2955
	6	2	0.5279	0.5942	0.6237	0.6357	0.6403	0.0215	0.1390	0.2090	0.2475	0.2674
		3	0.7275	0.7938	0.8232	0.8353	0.8399	0.1432	0.2606	0.3307	0.3692	0.3891
		5	0.8627	0.9290	0.9585	0.9705	0.9752	0.2525	0.3699	0.4400	0.4785	0.4983
	8	2	0.5219	0.6563	0.7419	0.7915	0.8183	0.4357	0.1488	0.2772	0.3743	0.4427
		3	0.6315	0.7659	0.8514	0.9010	0.9279	0.1007	0.2551	0.3835	0.4806	0.5490
		5	0.6742	0.8086	0.8941	0.9437	0.9705	0.1562	0.3106	0.4389	0.5361	0.6044
20	2	2	0.0609	0.0612	0.0612	0.0612	0.0612	0.0012	0.0024	0.0025	0.0026	0.0026
		3	0.1571	0.1571	0.1571	0.1571	0.1571	0.0062	0.0074	0.0076	0.0076	0.0076
		4	0.2921	0.2925	0.2925	0.2925	0.2925	0.0146	0.0158	0.0158	0.0160	0.0160
	5	3	0.4351	0.4393	0.4400	0.4402	0.4402	0.0446	0.0583	0.0619	0.0627	0.0629
		5	0.7902	0.7944	0.7952	0.7953	0.7954	0.1282	0.1419	0.1455	0.1463	0.1465
		7	0.9433	0.9475	0.9483	0.9484	0.9485	0.1851	0.1989	0.2035	0.2033	0.2035
	7	3	0.5882	0.6006	0.6035	0.6041	0.6043	0.0865	0.1218	0.1336	0.1372	0.1382
		5	0.8847	0.8970	0.9000	0.9006	0.9007	0.2058	0.2411	0.2529	0.2565	0.2575
		9	0.9818	0.9942	0.9971	0.9977	0.9979	0.2770	0.3123	0.3241	0.3277	0.3288
	12	5	0.8730	0.9437	0.9729	0.9836	0.9871	0.2707	0.4027	0.4792	0.5185	0.5369
		7	0.8836	0.9543	0.9835	0.9942	0.9977	0.2896	0.4216	0.4981	0.5374	0.5558
		10	0.8843	0.9550	0.9842	0.9949	0.9984	0.2919	0.4239	0.5004	0.5397	0.5581
	15	4	0.8749	0.9456	0.9747	0.9854	0.9890	0.1641	0.3349	0.4699	0.5639	0.6232
		8	0.8843	0.9550	0.9842	0.9949	0.9984	0.1836	0.3544	0.4894	0.5834	0.6427
		12	0.8843	0.9550	0.9842	0.9949	0.9985	0.1837	0.3545	0.4894	0.5835	0.6428
	18	5	0.4278	0.5981	0.7388	0.8414	0.9091	0.0376	0.1602	0.2960	0.4284	0.5450
		10	0.4279	0.5982	0.7389	0.8415	0.9093	0.0381	0.1608	0.2966	0.4289	0.5455
		17	0.4279	0.5982	0.7389	0.8415	0.9093	0.0381	0.1608	0.2966	0.4289	0.5455
30	10	2	0.3290	0.3388	0.3408	0.3412	0.3413	0.0170	0.0469	0.0558	0.0582	0.0588
		5	0.8889	0.8987	0.9007	0.9011	0.9011	0.2015	0.2314	0.2403	0.2427	0.2433
		7	0.9726	0.9824	0.9844	0.9848	0.9848	0.2598	0.2897	0.2986	0.3010	0.3016
	15	5	0.9200	0.9594	0.9718	0.9753	0.9761	0.2872	0.3792	0.4210	0.4377	0.4438
		10	0.9435	0.9829	0.9954	0.9988	0.9997	0.3243	0.4163	0.4581	0.4748	0.4809
		12	0.9435	0.9829	0.9954	0.9988	0.9997	0.3245	0.4165	0.4583	0.4750	0.4810
	20	8	0.8240	0.9243	0.9710	0.9900	0.9968	0.2520	0.4155	0.5239	0.5865	0.6188
		15	0.8240	0.9243	0.9710	0.9900	0.9969	0.2522	0.4156	0.5240	0.5867	0.6190
		18	0.8240	0.9243	0.9710	0.9900	0.9969	0.2522	0.4156	0.5240	0.5867	0.6190
	25	10	0.5611	0.7324	0.8524	0.9254	0.9651	0.0860	0.2460	0.4026	0.5357	0.6366
		15	0.5611	0.7324	0.8524	0.9254	0.9651	0.0860	0.2460	0.4026	0.5357	0.6366

Table 3: The values of $\alpha_{4(k)}\,(s,n;r,n)$ for $n = 10, 20, 30, 35$ and 40 and some selected values of s and r when $k = 2$ and $k = 3$

n	s	$k=2$					$k=3$				
		r					r				
		4	5	6	7	8	4	5	6	7	8
10	2	0.7280	0.6016	0.4758	0.3633	0.2696	0.8841	0.8211	0.7369	0.6432	0.5481
	4	0.7022	0.5934	0.4734	0.3626	0.2694	0.8128	0.7902	0.7245	0.6384	0.5464
	6	0.6169	0.5537	0.4567	0.3561	0.2670	0.6337	0.6785	0.6616	0.6056	0.5302
	8	0.4093	0.4141	0.3732	0.3102	0.2433	0.3332	0.4150	0.4564	0.4590	0.4321
20	2	0.8569	0.7563	0.6402	0.5212	0.4098	0.9639	0.9286	0.8756	0.8070	0.7265
	4	0.8547	0.7560	0.6401	0.5212	0.4098	0.9558	0.9268	0.8752	0.8069	0.7265
	6	0.8478	0.7545	0.6398	0.5211	0.4098	0.9331	0.9196	0.8732	0.8064	0.7264
	8	0.8315	0.7497	0.6386	0.5209	0.4097	0.8860	0.9003	0.8663	0.8041	0.7257
	10	0.7992	0.7379	0.6349	0.5158	0.4094	0.8057	0.8587	0.8475	0.7965	0.7229
	15	0.5808	0.6128	0.5725	0.4918	0.3978	0.4346	0.5691	0.6508	0.6761	0.6550
	18	0.2852	0.3546	0.3791	0.3627	0.3191	0.1463	0.2325	0.3151	0.3787	0.4149
30	2	0.9068	0.8272	0.7262	0.6135	0.4998	0.9829	0.9618	0.9268	0.8766	0.8122
	4	0.9063	0.8272	0.7262	0.6135	0.4998	0.9809	0.9615	0.9267	0.8766	0.8122
	6	0.9048	0.8270	0.7262	0.6135	0.4998	0.9753	0.9603	0.9265	0.8766	0.8122
	8	0.9013	0.8263	0.7261	0.6134	0.4997	0.9631	0.9570	0.9257	0.8764	0.8121
	10	0.8946	0.8248	0.7257	0.6134	0.4997	0.9413	0.9498	0.9236	0.8759	0.8120
	15	0.8504	0.8102	0.7216	0.6123	0.4995	0.8239	0.8945	0.9012	0.8677	0.8093
	20	0.7307	0.7515	0.6971	0.6033	0.4965	0.5902	0.7323	0.8056	0.8181	0.7860
	25	0.4680	0.5596	0.5786	0.5389	0.4649	0.2628	0.4014	0.5230	0.6059	0.6423
35	2	0.9215	0.8499	0.7556	0.6470	0.5341	0.9873	0.9704	0.9411	0.8976	0.8397
	4	0.9213	0.8499	0.7556	0.6470	0.5341	0.9862	0.9702	0.9411	0.8976	0.8397
	6	0.9204	0.8498	0.7556	0.6470	0.5341	0.9829	0.9696	0.9410	0.8976	0.8397
	8	0.9185	0.8495	0.7556	0.6470	0.5341	0.9758	0.9680	0.9407	0.8975	0.8397
	10	0.9148	0.8488	0.7555	0.6469	0.5341	0.9630	0.9644	0.9398	0.8973	0.8397
	15	0.8909	0.8422	0.7539	0.6466	0.5341	0.8920	0.9364	0.9303	0.8945	0.8389
	25	0.6871	0.7394	0.7093	0.6294	0.5280	0.5020	0.6622	0.7642	0.8047	0.7946
	30	0.4219	0.5289	0.5689	0.5476	0.4852	0.2107	0.3405	0.4663	0.5639	0.6202
40	2	0.9327	0.8679	0.7795	0.6750	0.5636	0.9998	0.9767	0.9515	0.9135	0.8613
	4	0.9323	0.8679	0.7795	0.6750	0.5636	0.9998	0.9767	0.9514	0.9135	0.8613
	6	0.9318	0.8678	0.7795	0.6750	0.5636	0.9998	0.9762	0.9514	0.9135	0.8613
	8	0.9308	0.8676	0.7795	0.6750	0.5636	0.9965	0.9753	0.9512	0.9135	0.8613
	10	0.9288	0.8672	0.7795	0.6750	0.5636	0.9887	0.9733	0.9508	0.9134	0.8612
	15	0.9142	0.8639	0.7788	0.6748	0.5636	0.9458	0.9582	0.9464	0.9122	0.8610
	20	0.8771	0.8516	0.7754	0.6740	0.5634	0.8442	0.9108	0.9273	0.9055	0.8588
	25	0.7965	0.8152	0.7615	0.6694	0.5621	0.6762	0.8014	0.8677	0.8772	0.8468
	30	0.6437	0.7216	0.7137	0.6482	0.5537	0.4360	0.5966	0.7189	0.7830	0.7937
	35	0.3823	0.4991	0.5557	0.5506	0.4999	0.1718	0.2916	0.4166	0.5231	0.5943

Table 4: Values of $\alpha_{7(k)}\,(s,t;1)$ for $1 \le s \le 7$ and $4 \le t \le 7$

s	$k=2$				$k=3$			
	t				t			
	4	5	6	7	4	5	6	7
1	0.1358	0.2016	0.2455	0.2748	0.0035	0.0527	0.1121	0.1565
2	0.3580	0.4239	0.4678	0.4970	0.1911	0.2402	0.2995	0.3440
3	0.5062	0.5720	0.6159	0.6452	0.3317	0.3808	0.4401	0.4846
4	0.6049	0.6707	0.7146	0.7339	0.4372	0.4862	0.5456	0.5901
5	0.6708	0.7366	0.7805	0.8098	0.5163	0.5654	0.6247	0.6692
6	0.7147	0.7805	0.8244	0.8537	0.5756	0.6247	0.6840	0.7285
7	0.7439	0.8098	0.8537	0.8829	0.6201	0.6692	0.7285	0.7730

Table 5: Values of $\alpha_{7(k)}\,(s,t;2)$ for $1 \le s \le 7$ and $4 \le t \le 7$

s	$k=2$				$k=3$			
	t				t			
	4	5	6	7	4	5	6	7
1	0.1502	0.2599	0.3477	0.4160	0.0011	0.0018	0.0759	0.1538
2	0.3684	0.4781	0.5659	0.6342	0.1126	0.2115	0.3005	0.3784
3	0.4605	0.5702	0.6580	0.7263	0.2262	0.3251	0.4141	0.4919
4	0.5019	0.6117	0.6994	0.7678	0.2873	0.3862	0.4752	0.5530
5	0.5213	0.6310	0.7188	0.7871	0.3213	0.4201	0.5091	0.5870
6	0.5304	0.6401	0.7279	0.7963	0.3405	0.4393	0.5284	0.6062
7	0.5349	0.6446	0.7323	0.8006	0.3516	0.4505	0.5395	0.6173

References

Ahmadi, J. and Balakrishnan, N. (2005). Distribution-free confidence intervals for quantile intervals based on current records. *Statistics and Probability Letters*, **75**, 190-202.

Ahmadi, J. and Balakrishnan, N. (2008). Nonparametric confidence intervals for quantile intervals and quantile differences based on record statistics. *Statistics and Probability Letters*, **78**, 1236-1245.

Ahmadi, J. and Balakrishnan, N. (2010). Prediction of order statistics and record values from two independent sequences. *Statistics*, **44**, 417-430.

Ahmadi, J., MirMostafaee, S. M. T. K. and Balakrishnan, N. (2010). Nonparametric prediction intervals for future record intervals based on order statistics. *Statistics and Probability Letters*, **80**, 1663-1672.

Ahmadi, J. and Balakrishnan, N. (2011). Distribution-free prediction intervals for order statistics based on record coverage. *Journal of the Korean Statistical Society*, **40**, 181-192.

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

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

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

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992). *A First Course in Order Statistics*. USA: Classics in Applied Mathematics.

Chacko, M. and Mary, M. S. (2013). Estimation and prediction based on k-record values from normal distribution. *Statistica*, **73**, 505-516.

Chacko, M. and Muraleedharan, L. (2018). Inference based on k-record values from generalized exponential distribution. *Statistica*, **78**, 37-56.

Chandler, K. (1952). The distribution and frequency of record values. *Journal of the Royal Statistical Society: Series B (Methodological)*, **14**, 220-228.

David, H. A. and Nagaraja, H.N. (2003). *Order Statistics*. Third Edition. New York: John Wiley & Sons.

Deheuvels, P. and Nevzorov, V. B. (1994). Limit laws for k-record times. *Journal of Statistical Planning and Inference*, **38**, 279-307.

Dziubdziela, W. and Kopociński, B. (1976). Limiting properties of the k-th record values. *Applicationes Mathematicae*, **2**, 187-190.

Glick, N. (1978). Breaking records and breaking boards. *The American Mathematical Monthly*, **85**, 2-26.

Guilbaud, O. (2004). Exact non-parametric confidence, prediction and tolerance intervals with progressive type-II censoring. *Scandinavian Journal of Statistics*, **31**, 265-281.

Gulati, S. and Padgett, W. J. (1994). Smooth nonparametric estimation of the distribution and density functions from record-breaking data. *Communications in Statistics-Theory and Methods*, **23**, 1259-1274.

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

Malinowska, I. and Szynal, D. (2004). On a family of Bayesian estimators and predictors for a Gumbel model based on the kth lower records. *Applicationes Mathematicae*, **1**, 107-115.

Minimol, S. and Thomas, P. Y. (2014). On characterization of Gompertz distribution by properties of generalized record values. *Journal of Statistical Theory and Applications*, **13**, 38-45.

Muraleedharan, L. and Chacko, M. (2019). Inference on Gompertz distribution based upper k- record values. *Journal of the Kerala Statistical Association,* **30**, 47-63.

Muraleedharan, L. and Chacko, M. (2022). Estimation and prediction based on k-record values from a generalized Pareto distribution. *International Journal of Mathematics and Statistics*, **23**, 59-75.

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

Paul, J. and Thomas, P. Y. (2015). On generalized upper (k) record values from Weibull distribution. *Statistica*, **75**, 313-330.