Journal of Statistical Research *https://doi.org/10.3329/jsr.v56i1.63946*
2022, Vol. 56, No. 1, pp. 55-73 *ISSN 0256 - 422 X*

INTERVAL PREDICTION OF ORDER STATISTICS AND RECORD VALUES USING CONCOMITANTS OF ORDER STATISTICS AND RECORD VALUES FOR MORGENSTERN FAMILY OF DISTRIBUTIONS

LAJI MURALEEDHARAN*
Department of Statistics, University of Kerala, Trivandrum 695581, India.
Email: lajikmr@gmail.com

MANOJ CHACKO
Department of Statistics, University of Kerala, Trivandrum 695581, India.
Email: manojchacko02@gmail.com

SUMMARY

In this paper, we discuss the problem of predicting intervals for future order statistics and k-record values based on observed concomitants of order statistics and observed concomitants of k-record values arising from a Morgenstern family of distributions. The coverage probabilities obtained are accurate and independent of the parent distribution. A real data set is also considered to exemplify the proposed methodologies developed in this paper.

Keywords and phrases: Order Statistics, k-record values, Concomitants of Order Statistics, Concomitants of k-record values, Morgenstern Family of Distributions, Prediction Intervals.

AMS Classification: Primary: 62G30; Secondary: 62E15

1 Introduction

Let $\{(X_i, Y_i), i = 1, 2, \ldots n\}$ be n independent and identically distributed (iid) bivariate random sample of observations arising from an absolutely continuous bivariate population with cumulative distribution function (cdf) $F(x, y)$ and joint probability density function (pdf) $f(x, y)$. Let $F_X(x)$ and $F_Y(y)$ be the marginal cdfs of X and Y, and let $f_X(x)$ and $f_Y(y)$ be the corresponding marginal pdfs of X and Y respectively. By arranging the X_i values in non-decreasing order of magnitude as $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$, the order statistics of the X variate will be obtained. Then the Y-variate associated with the rth order statistic $X_{r:n}$ is called the concomitant of $X_{r:n}$ and it is denoted by $Y_{[r:n]}$. The term concomitant of order statistic was first introduced by David (1973).

Concomitants of order statistics have found wide range of applications in the field of engineering, inference and prediction problems and double sampling plans. There are numerous studies available in the literature that deal with concomitants of order statistics. David et al. (1977) derived

* Corresponding author

the distribution of the rank of $Y_{[r:n]}$. The asymptotic behaviours of the rank of $Y_{[r:n]}$ were extensively discussed by David and Galambos (1974). Balasubramanian and Beg (1997, 1998) discussed the concomitant for Morgenstern type bivariate exponential distributions and Gumbel's bivariate exponential distributions, respectively and provided its recurrence relations for single and product moments. David and Nagaraja(1998) made a significant use of concomitants of order statistics in selection procedures when $k(< n)$ individuals are chosen on the basis of X values. Then the corresponding Y values represent the performance on an associated characteristic. The cdf and pdf of the concomitant of rth order statistic $Y_{[r:n]}$ are respectively given by (see, David, 1981)

$$F_{Y_{[r:n]}}(y) = \int_{-\infty}^{\infty} F_{Y|X}(y|x)\, f_{r:n}(x)\, dx \tag{1.1}$$

and

$$f_{Y_{[r:n]}}(y) = \int_{-\infty}^{\infty} f_{Y|X}(y|x)\, f_{r:n}(x)\, dx, \tag{1.2}$$

where $F_{Y|X}$ and $f_{Y|X}$ are respectively denote the conditional cdf and pdf of Y given X and $f_{r:n}(x)$ is the density function of $X_{r:n}$ which is given by (see, Arnold et al., 1992)

$$f_{r:n}(x) = \frac{1}{B(r, n-r+1)} [F_X(x)]^{r-1} [1 - F_X(x)]^{n-r} f_X(x), \ -\infty < x < \infty, \tag{1.3}$$

where $B(\cdot, \cdot)$ denotes the complete beta function.

Let $\{X_n, n \geq 1\}$ be a sequence of iid random variables with an absolutely continuous cdf $F_X(x)$ and pdf $f_X(x)$. 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 contained therein can be used to have a more in-depth look in this topic.

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 may be infinite. Such an issue does not arise if we use the k-records proposed by Dziubdziela and Kopocinski (1976). We use a 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 $n > 1$,

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

where $X_{r:m}$ denotes the rth order statistic in a sample of size m. Then the sequence of upper k-records $\left\{U_{n(k)}, n \geq 1\right\}$ 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

$$F_{n(k)}(x) = 1 - \left[\bar{F}_X(x)\right]^k \sum_{i=1}^{n-1} \frac{\left\{-k \log\left[\bar{F}_X(x)\right]\right\}^i}{i!}, \quad -\infty < x < \infty, \tag{1.4}$$

where $\bar{F} = 1 - F$. The pdf corresponds to the cdf (1.4) is given by

$$f_{n(k)}(x) = \frac{k^n}{\Gamma(n)} \left[-\log \bar{F}_X(x)\right]^{n-1} \left[\bar{F}_X(x)\right]^{k-1} f_X(x), \ -\infty < x < \infty, \tag{1.5}$$

where $\Gamma(\cdot)$ denotes the complete gamma function. The sequence of lower k-record values can be defined in a similar manner.

Let $\{(X_i, Y_i), i \geq 1\}$ be a sequence of iid bivariate random variables arising from a bivariate population with absolutely continuous cdf $F(x, y)$ and joint pdf $f(x, y)$. Let $\left\{U_{n(k)}, n \geq 1\right\}$ be the sequence of upper k-record values extracted from the X values. Then the Y-variable associated with the X-value which is quantified as the nth upper k-record value is called the concomitant of nth upper k-record value and is denoted by $U_{[n(k)]}$. An analogous definition deals with the concomitant of nth lower k-record value.

The cdf of the concomitant of nth upper k-record value $U_{[n(k)]}$ is defined as given below (see, Houchens, 1984).

$$F_{[n(k)]}(y) = \int_{-\infty}^{\infty} F_{Y|X}\left(y|x\right) f_{n(k)}(x) dx. \tag{1.6}$$

The pdf corresponds to the cdf (1.6) is obtained as

$$f_{[n(k)]}(y) = \int_{-\infty}^{\infty} f_{Y|X}\left(y|x\right) f_{n(k)}(x) dx, \tag{1.7}$$

where $f_{n(k)}$ is defined in (1.5).

In the area of modelling statistical data, families of distributions with members in a wide range of forms have aroused substantial interest. One standard method for solving modelling issues is choosing a family of distributions and selecting a member that best fits the observations. The most crucial factor in a modelling challenge is that the chosen family should be adaptable and including a wide range of models that can reflect any data scenario. In modelling bivariate data, when the prior information is in the form of marginal distributions, it is of advantage to consider families of bivariate distributions with specified marginals. Morgenstern families of distributions (MFD) is characterized by the specified marginal distribution functions $F_X(x)$ and $F_Y(y)$ of random variables X and Y respectively and a parameter α. A bivariate random variable (X, Y) whose distribution belongs to MFD if its cdf is given by (see, Kotz et al., 2000).

$$F\left(x, y\right) = F_X\left(x\right) F_Y\left(y\right) \left[1 + \alpha \left(1 - F_X\left(x\right)\right) \left(1 - F_Y\left(y\right)\right)\right], \ \ -1 \leq \alpha \leq 1. \tag{1.8}$$

The pdf corresponds to the cdf (1.8) is given by

$$f\left(x, y\right) = f_X\left(x\right) f_Y\left(y\right) \left[1 + \alpha \left(1 - 2F_X\left(x\right)\right) \left(1 - 2F_Y\left(y\right)\right)\right], \ \ -1 \leq \alpha \leq 1. \tag{1.9}$$

The estimation of parameters of Morgenstern type bivariate exponential distribution using concomitants of order statistics is extensively discussed by Chacko and Thomas (2011). Bairamov and Bekci(1999) looked at the concomitants for bivariate Farlie-Gumbel-Morgenstern type bivariate uniform distribution with uniform marginals by introducing additional parameters and found a recurrence relation between moments and moment generating function of concomitants order statistics.

In statistical inference, predicting future events based on current knowledge is a fundamental problem. It can be expressed in a variety of ways and 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.

There is a considerable amount of literature on the statistical prediction of future events. Several authors have considered prediction problems 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 sample are extensively discussed by Ahmadi and Balakrishnan (2005, 2008, 2010), Ahmadi et al. (2010) and Guilbaud (2004), respectively. However, to the best of our knowledge, the prediction of any future observations based on the observed sequence of concomitants of order statistics or concomitants of k-record values is not yet seen done in the available literature. Hence in this paper, based on the observed concomitants of order statistics and concomitants of k-record values arising from MFD with cdf given in (1.6), we obtain the two sample prediction intervals and the corresponding coverage probabilities for order statistics and k-record values from a future sample.

An explicit expression for the cdf and pdf of concomitants of order statistics and concomitants of k- record values is essential for finding the coverage probability of the prediction interval based on concomitants of order statistics and concomitants of k- record values. But the majority of well-known bivariate models, such as bivariate normal distribution, bivariate Pareto distribution, the pdfs and cdfs of concomitants of order statistics and concomitants of k- record values cannot be found explicitly. Suppose we have observed n concomitants of order statistics or concomitants of k-record values arising from MFD. Based on these data, we wish to construct the two sample prediction intervals for order statistics and k-record values from a future sample. Then the results developed in this paper can be used to find the prediction intervals and the corresponding prediction coefficients of order statistics and k-record values from a future sample.

The rest of this paper is structured as follows. In Section 2, we obtain the prediction intervals of future order statistics based on the observed sequence of concomitants of order statistics. In Section 3, we discuss the interval prediction of future order statistics based on the observed concomitants of k-record values. In Section 4, we obtain the interval prediction of future k-record values based on the observed concomitants of order statistics. The interval prediction of future record values based on the observed concomitants of k-record values are considered in Section 5. In Section 6, a real data set is used to exemplify the proposed methods developed in this paper and finally, some concluding remarks are made in Section 7.

2 Prediction Interval of Future Order Statistics Based on Concomitants of Order Statistics

In this section, we find the prediction intervals for future order statistics and derive the corresponding coverage probabilities based on the observed sequence of concomitants of order statistics arising from MFD. The cdf and pdf of the concomitant of rth order statistic $Y_{[r:n]}$ arising from MFD are respectively given by (see, Scaria and Nair, 1999)

$$F_{Y_{[r:n]}}(y) = F_Y(y)\left\{1+\alpha\left(\frac{n-2r+1}{n+1}\right)[1-F_Y(y)]\right\} \tag{2.1}$$

and

$$f_{Y_{[r:n]}}(y) = f_Y(y)\left\{1+\alpha\left(\frac{n-2r+1}{n+1}\right)[1-2F_Y(y)]\right\}. \tag{2.2}$$

Let $(X_i, Y_i), i = 1, 2, \ldots$ be a sequence of bivariate random sample of observations arising from a bivariate population with cdf $F(x, y)$. Let $\{T_n, n \geq 1\}$ be a sequence of observed concomitants of order statistics or concomitants of k-record values arising from $(X_i, Y_i), i = 1, 2, \ldots$. Suppose we are interested in obtaining an interval of the form (T_m, T_n), for $1 \leq m < n$, such that

$$\eta(m, n) = P(T_m \leq T \leq T_n) = 1 - \eta.$$

Then we refer to the interval (T_m, T_n) as a $100(1-\eta)\%$ prediction interval for the future observation T.

We can choose m and n so that $\eta(m, n)$ surpasses η_0 if the desired confidence level η_0 are supplied. Because $\eta(m, n)$ is a step function, the confidence coefficient may not equal η_0 but may be set to a value somewhat higher than η_0. Furthermore, the choice of m and n is not unique. We would like to generate a prediction interval as short as possible among all prediction intervals with the same level for a given confidence level of η_0. First, notice that the two-sided prediction intervals exist for a given η_0, if and only if, for large m,

$$P(T_1 \leq T \leq T_m) \geq \eta_0.$$

The following theorem establishes the prediction intervals and the corresponding coverage probabilities of order statistics from a future sample based on the observed concomitants of order statistics.

Theorem 2.1. Let $\{Y_{[r:n]}, r = 1, 2, \ldots, n\}$ be n observed concomitants of order statistics arising from a MFD with cdf given in (1.8). Let F_Y be the marginal cdf of Y and let $Y_{1:m} \leq Y_{2:m} \leq \cdots \leq Y_{m:m}$ be the order statistics of a future random sample of size m arising from the same cdf F_Y. Then $(Y_{[s:n]}, Y_{[t:n]})$, for $1 \leq s < t \leq n$, is a prediction interval for the rth order statistic $Y_{r:m}$, for $1 \leq r \leq m$, with the corresponding prediction coefficient, being free of F_Y and is given by

1. If $Y_{[s:n]} < Y_{[t:n]}$ and $0 < \alpha < 1$, then $(Y_{[s:n]}, Y_{[t:n]})$ is a prediction interval with the corresponding prediction coefficient given by

$$\eta_1(s, t, r; \alpha, m, n) = \frac{2\alpha r(t-s)(m-r+1)}{(n+1)(m+1)(m+2)}. \tag{2.3}$$

2. If $Y_{[t:n]} < Y_{[s:n]}$ and $-1 < \alpha < 0$, then $\left(Y_{[t:n]}, Y_{[s:n]}\right)$ is a prediction interval with the corresponding prediction coefficient given by

$$\eta_2\left(s,t,r;\alpha,m,n\right) = \frac{2\alpha r\left(s-t\right)\left(m-r+1\right)}{\left(n+1\right)\left(m+1\right)\left(m+2\right)}. \tag{2.4}$$

Proof. First we consider the case when $Y_{[s:n]} < Y_{[t:n]}$. Then for any fixed real number v and $1 \leq s < t \leq n$, then we have

$$\begin{aligned} P\left(Y_{[s:n]} \leq v\right) &= P\left(Y_{[s:n]} \leq v, Y_{[t:n]} < v\right) + P\left(Y_{[s:n]} \leq v, Y_{[t:n]} \geq v\right) \\ &= P\left(Y_{[t:n]} < v\right) + P\left(Y_{[s:n]} \leq v \leq Y_{[t:n]}\right). \end{aligned}$$

Hence

$$P\left(Y_{[s:n]} \leq v \leq Y_{[t:n]}\right) = P\left(Y_{[s:n]} \leq v\right) - P\left(Y_{[t:n]} \leq v\right). \tag{2.5}$$

Using (2.1), (2.5) can be expressed as

$$P\left(Y_{[s:n]} \leq v \leq Y_{[t:n]}\right) = \frac{2\alpha\left(t-s\right)}{\left(n+1\right)} F_Y\left(v\right)\left[1 - F_Y\left(v\right)\right]. \tag{2.6}$$

Now for $s < t$, and using the conditioning arguments, we can write

$$\begin{aligned} P\left(Y_{[s:n]} \leq Y_{r:m} \leq Y_{[t:n]}\right) &= \int_{-\infty}^{\infty} P\left(Y_{[s:n]} \leq Y_{r:m} \leq Y_{[t:n]} | Y_{r:m} = v\right) f_{r:m}\left(v\right) dv \\ &= \int_{-\infty}^{\infty} P\left(Y_{[s:n]} \leq v \leq Y_{[t:n]}\right) f_{r:m}\left(v\right) dv \\ &= \frac{2\alpha\left(t-s\right)}{\left(n+1\right)} \int_{-\infty}^{\infty} F_Y\left(v\right)\left[1 - F_Y\left(v\right)\right] f_{r:m}\left(v\right) dv \\ &= \frac{2\alpha r\left(t-s\right)\left(m-r+1\right)}{\left(n+1\right)\left(m+1\right)\left(m+2\right)}. \end{aligned} \tag{2.7}$$

Thus for $0 < \alpha < 1$ and $Y_{[s:n]} < Y_{[t:n]}$, we have

$$\eta_1\left(s,t,r;\alpha,m,n\right) = P\left(Y_{[s:n]} \leq Y \leq Y_{[s:n]}\right) = \frac{2\alpha r\left(t-s\right)\left(m-r+1\right)}{\left(n+1\right)\left(m+1\right)\left(m+2\right)}.$$

By a similar arguments, the result follows for the case when $-1 < \alpha < 0$ and $Y_{[t:n]} < Y_{[s:n]}$. Hence the proof. □

We have evaluated the coverage probabilities $\eta_1\left(s,t,r;\alpha,m,n\right)$ for different values of $s-t$, r and α for $n = 20, 30$ and $m = 15, 25$. The values are presented in Table 1. It can be observed that coverage probabilities improve with the increase of α. When $-1 < \alpha < 0$, we can write $\eta_2\left(s,t,r;\alpha,m,n\right) = \eta_1\left(s,t,r;-\alpha,m,n\right)$, hence one can use Table 1 for evaluating (2.4).

Remark 2.1. When the parameters are presented in the coverage probabilities of the prediction intervals of future event, Escobar and Meeker (1999), suggests different calibration methods for computing the probabilities of the prediction intervals. In this approach, maximum likelihood estimate (MLE) for α can be used to predict a future independent observation from the observed data. They have discussed in detail about the problem of prediction in case of log-location-scale distributions such as Weibull or lognormal distributions. Cox (1975) suggested a large sample approximate method based on MLEs, that can be used to calibrate or correct a naive prediction coefficient. According to this approach, to calibrate the prediction coefficient by evaluating the value $1-\eta_c$ such that

$$1-\eta_c = P\left(T_1 \leq T \leq T_2 | \hat{\alpha}\right) = P\left(\hat{t}_{\frac{\eta_c}{2}} \leq T \leq \hat{t}_{1-\frac{\eta_c}{2}} | \hat{\alpha}\right),$$

where $\hat{t}_p$ is the MLE of the pth quantile of T.

3 Prediction Interval of Future Order Statistics Based on Concomitants of *k*-Record Values

In this section, we find the prediction intervals for future order statistics and obtain the corresponding coverage probabilities based on the observed sequence of concomitants of k-record values.

By Chacko and Mary (2013), the cdf of nth concomitant of upper k-record value $U_{[n(k)]}$ arising from MFD is given by

$$F_{[n(k)]}(y) = F_Y(y)\left\{1+\alpha\left[1-2\left(\frac{k}{1+k}\right)^n\right]\left[F_Y(y)-1\right]\right\}. \tag{3.1}$$

The pdf corresponds to the cdf (3.1) is obtained as

$$f_{[n(k)]}(y) = f_Y(y)\left\{1+\alpha\left[1-2\left(\frac{k}{1+k}\right)^n\right]\left[2F_Y(y)-1\right]\right\}. \tag{3.2}$$

The cdf and pdf of the concomitant of nth lower k-record value $L_{[n(k)]}$ arising from MFD are respectively given by

$$F^*_{[n(k)]}(y) = F_Y(y)\left\{1+\alpha\left[2\left(\frac{k}{1+k}\right)^n-1\right]\left[F_Y(y)-1\right]\right\} \tag{3.3}$$

and

$$f^*_{[n(k)]}(y) = f_Y(y)\left\{1+\alpha\left[2\left(\frac{k}{1+k}\right)^n-1\right]\left[2F_Y(y)-1\right]\right\}. \tag{3.4}$$

The following theorem provides the interval prediction and exact expression for the coverage probabilities of future order statistics based on the observed sequence of concomitants of upper k- record values.

Theorem 3.1. Let $\left\{U_{[n(k)]}, n \geq 1\right\}$ be a sequence of observed concomitants of upper k-record values arising from a MFD with cdf given in (1.8). Let $Y_{1:m} \leq Y_{2:m} \leq \cdots \leq Y_{m:m}$ be the order

Table 1: Values of $\eta_1\,(s,t,r;\alpha,m,n)$ for some selected values of $t-s, r, n, m$ and α.

n	m	r	$t-s$	α					
				0.30	0.50	0.75	0.90	0.95	1.00
20	15	8	6	0.04034	0.06723	0.10084	0.12101	0.12773	0.13445
			8	0.05378	0.08964	0.13445	0.16134	0.17031	0.17927
			10	0.06723	0.11204	0.16807	0.20168	0.21289	0.22410
			12	0.08067	0.13445	0.20168	0.24202	0.25546	0.26891
			15	0.10084	0.16807	0.25210	0.30252	0.31933	0.33614
			18	0.12101	0.20168	0.30252	0.36303	0.38319	0.40336
		12	6	0.03025	0.05042	0.07563	0.09076	0.09580	0.10084
			8	0.04034	0.06723	0.10084	0.12101	0.12773	0.13445
			10	0.05042	0.08403	0.12605	0.15126	0.15966	0.16807
			12	0.06050	0.10084	0.15126	0.18151	0.19160	0.20168
			15	0.07563	0.12605	0.18908	0.22689	0.23950	0.25210
			18	0.09076	0.15126	0.22689	0.27227	0.28739	0.30252
30	25	12	6	0.02779	0.04632	0.06948	0.08337	0.08801	0.09264
			8	0.03706	0.06176	0.09264	0.11117	0.11734	0.12352
			10	0.04632	0.07720	0.11580	0.13896	0.14668	0.15440
			12	0.05558	0.09264	0.13896	0.16675	0.17601	0.18528
			15	0.06948	0.11580	0.17370	0.20844	0.22002	0.23160
			18	0.08337	0.13896	0.20844	0.25012	0.26402	0.27792
			20	0.09264	0.15440	0.23160	0.27792	0.29336	0.30880
			25	0.11580	0.19300	0.28950	0.34739	0.36669	0.38599
		15	6	0.02730	0.04549	0.06824	0.08189	0.08644	0.09098
			8	0.03639	0.06066	0.09098	0.10918	0.11525	0.12131
			10	0.04549	0.07582	0.11373	0.13648	0.14406	0.15164
			12	0.05459	0.09098	0.13648	0.16377	0.17287	0.18197
			15	0.06824	0.11373	0.17060	0.20471	0.21609	0.22746
			18	0.08189	0.13648	0.20471	0.24566	0.25931	0.27295
			20	0.09098	0.15164	0.22746	0.27295	0.28812	0.30328
			25	0.11373	0.18955	0.28433	0.34119	0.36015	0.37910
		20	6	0.01985	0.03309	0.04963	0.05955	0.06286	0.06617
			8	0.02647	0.04411	0.06617	0.07940	0.08382	0.08823
			10	0.03309	0.05514	0.08271	0.09926	0.10477	0.11028
			12	0.03970	0.06617	0.09926	0.11911	0.12572	0.13234
			15	0.04963	0.08271	0.12407	0.14888	0.15715	0.16543
			18	0.05955	0.09926	0.14888	0.17866	0.18859	0.19851
			20	0.06617	0.11028	0.16543	0.19851	0.20954	0.22057
			25	0.08271	0.13785	0.20678	0.24814	0.26192	0.27571

statistics from a future random sample of size m arising from the same cdf F_Y. Then for $1 \leq s < t$, the prediction interval for the rth future order statistic $Y_{r:m}$, for $1 \leq r \leq m$, and the corresponding prediction coefficient, being free of F_Y, are given below.

1. If $U_{[s(k)]} < U_{[t(k)]}$ and $0 < \alpha < 1$, then the prediction interval is $\left(U_{[s(k)]}, U_{[t(k)]}\right)$ with corresponding prediction coefficient

$$\eta_{3(k)}\left(s,t,r;\alpha,m\right) = 2\alpha\left\{\left(\frac{k}{1+k}\right)^{s} - \left(\frac{k}{1+k}\right)^{t}\right\}\frac{r\left(m-r+1\right)}{\left(m+1\right)\left(m+2\right)}. \tag{3.5}$$

2. If $U_{[t(k)]} < U_{[s(k)]}$ and $-1 < \alpha < 0$, then the prediction interval is $\left(U_{[t(k)]}, U_{[s(k)]}\right)$ with corresponding prediction coefficient

$$\eta_{4(k)}\left(s,t,r;\alpha,m\right) = 2\alpha\left\{\left(\frac{k}{1+k}\right)^{t} - \left(\frac{k}{1+k}\right)^{s}\right\}\frac{r\left(m-r+1\right)}{\left(m+1\right)\left(m+2\right)}. \tag{3.6}$$

Proof. For any fixed real number v, suppose $U_{[s(k)]} < U_{[t(k)]}$ for $1 \leq s < t$ and $0 < \alpha < 1$, then 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

$$P\left(U_{[s(k)]} \leq v \leq U_{[t(k)]}\right) = P\left(U_{[s(k)]} \leq v\right) - P\left(U_{[t(k)]} \leq v\right). \tag{3.7}$$

Using (3.1), (3.7) can be expressed as

$$P\left(U_{[s(k)]} \leq v \leq U_{[t(k)]}\right) = 2\alpha\left\{\left(\frac{k}{1+k}\right)^{t} - \left(\frac{k}{1+k}\right)^{s}\right\}F_Y\left(v\right)\left[F_Y\left(v\right) - 1\right]. \tag{3.8}$$

Now for $s < t$, and using the conditioning arguments, we can write

$$\begin{aligned} P\left(U_{[s(k)]} \leq Y_{r:m} \leq U_{[t(k)]}\right) &= \int_{-\infty}^{\infty} P\left(U_{[s(k)]} \leq Y_{r:m} \leq U_{[t(k)]} | Y_{r:m} = v\right) f_Y\left(v\right) dv \\ &= \int_{-\infty}^{\infty} P\left(R_{[s(k)]} \leq v \leq R_{[t(k)]}\right) f_{r:m}\left(v\right) dv \\ &= 2\alpha\left\{\left(\frac{k}{1+k}\right)^{t} - \left(\frac{k}{1+k}\right)^{s}\right\}\int_{-\infty}^{\infty} F_Y\left(v\right)\left[F_Y\left(v\right) - 1\right] f_{r:m}\left(v\right) dv \\ &= 2\alpha\left\{\left(\frac{k}{1+k}\right)^{s} - \left(\frac{k}{1+k}\right)^{t}\right\}\frac{r\left(m-r+1\right)}{\left(m+1\right)\left(m+2\right)}. \end{aligned}$$

Thus for $0<\alpha<1$ and $U_{[s(k)]}<U_{[t(k)]}$, we have

$$\begin{aligned}\eta_{3(k)}\left(s,t,r;\alpha,m\right)&=P\left(U_{[s(k)]}\leq Y\leq U_{[t(k)]}\right)\\&=2\alpha\left\{\left(\frac{k}{1+k}\right)^{s}-\left(\frac{k}{1+k}\right)^{t}\right\}\frac{r\left(m-r+1\right)}{\left(m+1\right)\left(m+2\right)}.\end{aligned}$$

By a similar arguments, the result follows for the case when $-1<\alpha<0$ and $U_{[t(k)]}<U_{[s(k)]}$. Hence the proof. □

We have evaluated the coverage probabilities $\eta_{3(k)}\left(s,t,r;\alpha,m\right)$ under $m=20$, $k=1,2,3$ and $\alpha=0.5,0.75,0.9,0.95$ for $r=5,8,10$ and 15. The values are presented in Table 2. It can be observed that coverage probabilities improve with the increase of α and k. When $-1<\alpha<0$, we can write $\eta_{4(k)}\left(s,t,r;\alpha,m\right)=\eta_{3(k)}\left(s,t,r;-\alpha,m\right)$, hence one can use Table 2 for evaluating $\eta_{4(k)}\left(s,t,r;\alpha,m\right)$.

Now we consider the following theorem which establishes the prediction intervals and corresponding prediction coefficients of order statistics from a future sample based on the observed sequence of concomitants of lower k-record values.

Theorem 3.2. Let $\left\{L_{[n(k)]},n\geq1\right\}$ be a sequence of observed concomitants of lower k-record values arising from a MFD with cdf given in (1.8). Let $Y_{1:m}\leq Y_{2:m}\leq\cdots\leq Y_{m:m}$ be the order statistics from a future random sample of size m arising from the same cdf F_Y. Then for $1\leq s<t$, the prediction interval for future rth order statistics $Y_{r:m}$, for $1\leq r\leq m$, and the corresponding prediction coefficient, being free of F_Y, are given below

1. If $L_{[s(k)]}<L_{[t(k)]}$ and $-1<\alpha<0$ then the prediction interval is $\left(L_{[s(k)]},L_{[t(k)]}\right)$ with corresponding prediction coefficient

$$\eta_{5(k)}\left(s,t,r;\alpha,m\right)=2\alpha\left\{\left(\frac{k}{1+k}\right)^{t}-\left(\frac{k}{1+k}\right)^{s}\right\}\frac{r\left(m-r+1\right)}{\left(m+1\right)\left(m+2\right)}.\tag{3.9}$$

2. If $L_{[t(k)]}<L_{[s(k)]}$ and $0<\alpha<1$ and then the prediction interval is $\left(L_{[t(k)]},L_{[s(k)]}\right)$ with corresponding prediction coefficient

$$\eta_{6(k)}\left(s,t,r;\alpha,m\right)=2\alpha\left\{\left(\frac{k}{1+k}\right)^{s}-\left(\frac{k}{1+k}\right)^{t}\right\}\frac{r\left(m-r+1\right)}{\left(m+1\right)\left(m+2\right)}.\tag{3.10}$$

Proof. The proof of the theorem directly follows from Theorem 3.1 and thus omitted. □

Remark 3.1. In the light of Theorem 3.1 and Theorem 3.2, we can observe the following two identities.

1. $\eta_{5(k)}\left(s,t,r;\alpha,m\right)=\eta_{4(k)}\left(s,t,r;\alpha,m\right).$
2. $\eta_{6(k)}\left(s,t,r;\alpha,m\right)=\eta_{3(k)}\left(s,t,r;\alpha,m\right).$

Thus one can evaluate $\eta_{5(k)}\left(s,t,r;\alpha,m\right)$ and $\eta_{6(k)}\left(s,t,r;\alpha,m\right)$ by using Table 2.

4 Prediction Interval of Future *k*-Record Values Based on Concomitants of Order Statistics

In this section, we find the prediction intervals for the future k-record values and derive the corresponding coverage probabilities based on the observed sequence of concomitants of order statistics. The following theorem establishes the interval prediction and the corresponding prediction coefficient of rth future (upper or lower) k-record values.

Theorem 4.1. Let $\left\{Y_{[i:n]}, i=1,2,\ldots,n\right\}$ be n observed concomitants of order statistics arising from a MFD with cdf given in (1.8). Then for $1 \le s < t \le n$, the prediction interval for rth future k-record value $R_{r(k)}$ and the corresponding prediction coefficient, being free of F_Y, are given below.

1. If $0 < \alpha < 1$ and $Y_{[s:n]} < Y_{[t:n]}$, then $\left(Y_{[s:n]}, Y_{[t:n]}\right)$ is a prediction interval with the corresponding prediction coefficient given by

$$\eta_{7(k)}\left(s,t,r;\alpha,n\right) = \frac{2\alpha\left(t-s\right)}{\left(n+1\right)}\left[\left(\frac{k}{k+1}\right)^r - \left(\frac{k}{k+2}\right)^r\right]. \tag{4.1}$$

2. If $-1 < \alpha < 0$ and $Y_{[t:n]} < Y_{[s:n]}$, then $\left(Y_{[t:n]}, Y_{[s:n]}\right)$ is a prediction interval with the corresponding prediction coefficient given by

$$\eta_{8(k)}\left(s,t,r;\alpha,n\right) = \frac{2\alpha\left(s-t\right)}{\left(n+1\right)}\left[\left(\frac{k}{k+1}\right)^r - \left(\frac{k}{k+2}\right)^r\right]. \tag{4.2}$$

Proof. Let $R_{r(k)}$ be the rth future upper k-record value with pdf given in (1.5).
First we consider the case when $Y_{[s:n]} < Y_{[t:n]}$. Then for any fixed real number v and $1 \le s < t \le n$,

$$P\left(Y_{[s:n]} \le v \le Y_{[t:n]}\right) = \frac{2\alpha\left(t-s\right)}{\left(n+1\right)} F_Y\left(v\right)\left[1 - F_Y\left(v\right)\right]. \tag{4.3}$$

Now for $s < t$, and using the conditioning arguments, we can write

$$\begin{aligned} P\left(Y_{[s:n]} \le R_{r(k)} \le Y_{[t:n]}\right) &= \int_{-\infty}^{\infty} P\left(Y_{[s:n]} \le R_{r(k)} \le Y_{[t:n]} | R_{r(k)} = v\right) f_Y\left(v\right) dv \\ &= \int_{-\infty}^{\infty} P\left(Y_{[s:n]} \le v \le Y_{[t:n]}\right) f_{r(k)}\left(v\right) dv \\ &= \frac{2\alpha\left(t-s\right)}{\left(n+1\right)} \int_{-\infty}^{\infty} F_Y\left(v\right)\left[1 - F_Y\left(v\right)\right] f_{r(k)}\left(v\right) dv. \end{aligned} \tag{4.4}$$

Table 2: Values of $\eta_{3(k)}\,(s,t,r;\alpha,m)$ for some selected values of s,t,r,m and α

m	α	r	$k=1$					$k=2$					$k=3$				
			(s,t)					(s,t)					(s,t)				
			$(1,8)$	$(2,10)$	$(1,12)$	$(2,15)$	$(2,18)$	$(1,8)$	$(2,10)$	$(1,12)$	$(2,15)$	$(2,18)$	$(1,8)$	$(2,10)$	$(1,12)$	$(2,15)$	$(2,18)$
20	0.50	5	0.08590	0.08641	0.08654	0.08657	0.08658	0.10868	0.11244	0.11411	0.11504	0.11532	0.11253	0.12012	0.12439	0.12756	0.12889
	0.75		0.12886	0.12962	0.12981	0.12986	0.12987	0.16303	0.16866	0.17116	0.17257	0.17298	0.16880	0.18018	0.18658	0.19133	0.19334
	0.90		0.15463	0.15554	0.15577	0.15583	0.15584	0.19563	0.20239	0.20539	0.20708	0.20758	0.20256	0.21621	0.22389	0.22960	0.23201
	0.95		0.16322	0.16418	0.16442	0.16449	0.16450	0.20650	0.21363	0.21680	0.21858	0.21911	0.21382	0.22823	0.23633	0.24236	0.24490
	1.00		0.17181	0.17282	0.17308	0.17315	0.17316	0.21737	0.22487	0.22821	0.23009	0.23065	0.22507	0.24024	0.24877	0.25511	0.25779
	0.50	8	0.11167	0.11233	0.11250	0.11255	0.11255	0.14129	0.14617	0.14834	0.14956	0.14992	0.14629	0.15615	0.16170	0.16582	0.16756
	0.75		0.16751	0.16850	0.16875	0.16882	0.16883	0.21193	0.21925	0.22251	0.22434	0.22488	0.21944	0.23423	0.24255	0.24873	0.25134
	0.90		0.20101	0.20220	0.20250	0.20259	0.20260	0.25432	0.26310	0.26701	0.26920	0.26986	0.26333	0.28108	0.29106	0.29848	0.30161
	0.95		0.21218	0.21344	0.21375	0.21384	0.21385	0.26845	0.27772	0.28184	0.28416	0.28485	0.27796	0.29669	0.30723	0.31506	0.31837
	1.00		0.22335	0.22467	0.22500	0.22509	0.22511	0.28258	0.29234	0.29667	0.29912	0.29984	0.29259	0.31231	0.32340	0.33165	0.33512
	0.50	10	0.11812	0.11882	0.11899	0.11904	0.11905	0.14944	0.15460	0.15690	0.15819	0.15857	0.15474	0.16516	0.17103	0.17539	0.17723
	0.75		0.17718	0.17822	0.17848	0.17856	0.17857	0.22416	0.23190	0.23534	0.23728	0.23785	0.23210	0.24775	0.25654	0.26308	0.26584
	0.90		0.21261	0.21387	0.21418	0.21427	0.21428	0.26899	0.27828	0.28241	0.28474	0.28542	0.27852	0.29729	0.30785	0.31570	0.31901
	0.95		0.22442	0.22575	0.22608	0.22618	0.22619	0.28394	0.29374	0.29810	0.30055	0.30128	0.29400	0.31381	0.32496	0.33324	0.33674
	1.00		0.23624	0.23763	0.23798	0.23808	0.23809	0.30000	0.30920	0.31379	0.31637	0.31714	0.30947	0.33033	0.34206	0.35078	0.35446
	0.50	15	0.09664	0.09721	0.09736	0.09740	0.09740	0.12227	0.12649	0.12837	0.12943	0.12974	0.12660	0.13513	0.13993	0.14350	0.14501
	0.75		0.14496	0.14582	0.14603	0.14610	0.14610	0.18340	0.18974	0.19255	0.19414	0.19461	0.18990	0.20270	0.20990	0.21525	0.21751
	0.90		0.17396	0.17498	0.17524	0.17531	0.17532	0.22008	0.22769	0.23106	0.23297	0.23353	0.22788	0.24324	0.25188	0.25830	0.26101
	0.95		0.18362	0.18470	0.18497	0.18505	0.18506	0.23231	0.24033	0.24390	0.24591	0.24650	0.24054	0.25675	0.26587	0.27265	0.27551
	1.00		0.19328	0.19442	0.19471	0.19479	0.19480	0.24454	0.25298	0.25674	0.25885	0.25948	0.25320	0.27027	0.27987	0.28700	0.29001

By using (1.5) in (4.4), we obtain the following

$$P\big(Y_{[s:n]} \le R_{r(k)} \le Y_{[t:n]}\big) = \frac{2\alpha k^r\,(t-s)}{\Gamma(r)\,(n+1)} \times \int_{-\infty}^{\infty} \left[-\log\left(1-F_Y(v)\right)\right]^{r-1} \left[1-F_Y(v)\right]^k F_Y(v) f_Y(v)dv \quad (4.5)$$

Taking $-\log\left(1-F_Y(v)\right) = u$ and evaluating the integral of (4.5), we finally arrive at

$$\eta_{7(k)}\left(s,t,r;\alpha,n\right) = \frac{2\alpha\,(t-s)}{(n+1)} \left[\left(\frac{k}{k+1}\right)^r - \left(\frac{k}{k+2}\right)^r\right]. \quad (4.6)$$

By a similar arguments, the result follows for the case when $Y_{[t:n]} < Y_{[s:n]}$ and $-1 < \alpha < 0$. Hence the proof. □

We have evaluated the coverage probabilities $\eta_{7(k)}\left(s,t,r;\alpha,n\right)$ under $n = 10, 12$, various values of α , $t-s$ and r. The values are presented in Table 3. It can be observed that coverage probabilities improve with the increase of α and k. Notice that when $-1 < \alpha < 0$, $\eta_{8(k)}\left(s,t,r;\alpha,n\right) = \eta_{7(k)}\left(s,t,r;-\alpha,n\right)$. Thus one can use Table 3 for evaluating $\eta_{8(k)}\left(s,t,r;\alpha,n\right)$.

5 Prediction Interval of Future Record Values Based on the Concomitants of *k*-Record Values

In this section, we find the prediction intervals for the future record values and obtain the corresponding coverage probabilities based on the observed sequence of concomitants of *k*- record values. The following theorem establishes the interval prediction and the corresponding coverage probabilities of upper record values from a future sample based on the observed sequence of concomitants of upper *k*-record values. One can observe that the same results follow the interval prediction and the corresponding coverage probabilities of lower record values from a future sample based on the observed sequence of concomitants of lower *k*-record values.

Table 3: Values of $\eta_{7(k)}\left(s,t,r;\alpha,n\right)$ for some selected values of s,t,r,n and α.

n	r	$t-s$	$k=1$				$k=2$				$k=3$			
			0.75	0.90	0.95	1.00	0.75	0.90	0.95	1.00	0.75	0.90	0.95	1.00
10	3	6	0.07197	0.08636	0.09116	0.09596	0.14015	0.16818	0.17753	0.18687	0.16844	0.20213	0.21336	0.22459
		8	0.09596	0.11515	0.12155	0.12795	0.18687	0.22424	0.23670	0.24916	0.22459	0.26951	0.28448	0.29945
		9	0.10795	0.12955	0.13674	0.14394	0.21023	0.25227	0.26629	0.28030	0.25266	0.30320	0.32004	0.33689
	5	6	0.02220	0.02664	0.02812	0.02960	0.08218	0.09861	0.10409	0.10957	0.13054	0.15664	0.16535	0.17405
		8	0.02960	0.03552	0.03750	0.03947	0.10957	0.13148	0.13879	0.14609	0.17405	0.20886	0.22046	0.23207
		9	0.03330	0.03996	0.04218	0.04440	0.12326	0.14792	0.15613	0.16435	0.19580	0.23497	0.24802	0.26107
12	4	6	0.03472	0.04167	0.04398	0.04630	0.09348	0.11218	0.11841	0.12464	0.12933	0.15519	0.16381	0.17244
		8	0.04630	0.05556	0.05864	0.06173	0.12464	0.14957	0.15788	0.16619	0.17244	0.20692	0.21842	0.22992
		10	0.05787	0.06944	0.07330	0.07716	0.15580	0.18697	0.19735	0.20774	0.21555	0.25865	0.27302	0.28739
	6	6	0.00987	0.01184	0.01250	0.01316	0.04996	0.05995	0.06328	0.06662	0.09092	0.10910	0.11516	0.12122
		8	0.01316	0.01579	0.01667	0.01754	0.06662	0.07994	0.08438	0.08882	0.12122	0.14546	0.15355	0.16163
		10	0.01645	0.01974	0.02083	0.02193	0.08327	0.09992	0.10547	0.11103	0.15153	0.18183	0.19193	0.20203

Theorem 5.1. Let $\left\{U_{[n(k)]}, n \geq 1\right\}$ be a sequence of observed concomitants of upper k-record values arising from a MFD with cdf given in (1.8). Then for $1 \leq s < t$, the prediction interval for rth future upper record value U_r and the corresponding prediction coefficient, being free of F_Y, are given below

1. If $0 < \alpha < 1$ and $U_{[s(k)]} < U_{[t(k)]}$, then the prediction interval is $\left(U_{[s(k)]}, U_{[t(k)]}\right)$ and the corresponding prediction coefficient is given by

$$\eta_{9(k)}\left(s,t,r;\alpha\right) = 2\alpha\left[\left(\frac{k}{1+k}\right)^{s} - \left(\frac{k}{1+k}\right)^{t}\right]\left(\frac{1}{2^{r}} - \frac{1}{3^{r}}\right). \tag{5.1}$$

2. If $-1 < \alpha < 0$ and $U_{[t(k)]} < U_{[s(k)]}$ then the prediction interval is $\left(U_{[t(k)]}, U_{[s(k)]}\right)$ and the corresponding prediction coefficient is given by

$$\eta_{10(k)}\left(s,t,r;\alpha\right) = 2\alpha\left[\left(\frac{k}{1+k}\right)^{t} - \left(\frac{k}{1+k}\right)^{s}\right]\left(\frac{1}{2^{r}} - \frac{1}{3^{r}}\right). \tag{5.2}$$

Proof. For any fixed real number v, suppose $U_{[s(k)]} < U_{[t(k)]}$ for $1 \leq s < t$, we have obtained

$$P\left(U_{[s(k)]} \leq v \leq U_{[t(k)]}\right) = 2\alpha\left\{\left(\frac{k}{1+k}\right)^{s} - \left(\frac{k}{1+k}\right)^{t}\right\} F_Y\left(v\right)\left[F_Y\left(v\right) - 1\right]. \tag{5.3}$$

Now for $s < t$, and using the conditioning arguments, we can write

$$\begin{aligned} P\left(U_{[s(k)]} \leq U_r \leq U_{[t(k)]}\right) &= \int_{-\infty}^{\infty} P\left(U_{[s(k)]} \leq U_r \leq U_{[t(k)]} | U_r = v\right) f_Y\left(v\right) dv \\ &= \int_{-\infty}^{\infty} P\left(U_{[s(k)]} \leq v \leq U_{[t(k)]}\right) f_r\left(v\right) dv \\ &= 2\alpha\left[\left(\frac{k}{1+k}\right)^{s} - \left(\frac{k}{1+k}\right)^{t}\right] \int_{-\infty}^{\infty} F_Y\left(v\right)\left[1 - F_Y\left(v\right)\right] f_r\left(v\right) dv \\ &= 2\alpha\left[\left(\frac{k}{1+k}\right)^{s} - \left(\frac{k}{1+k}\right)^{t}\right]\left(\frac{1}{2^{r}} - \frac{1}{3^{r}}\right). \end{aligned}$$

Thus for $U_{[s(k)]} < U_{[t(k)]}$ and $0 < \alpha < 1$, we have

$$\eta_{9(k)}\left(s,t,r;\alpha\right) = 2\alpha\left[\left(\frac{k}{1+k}\right)^{s} - \left(\frac{k}{1+k}\right)^{t}\right]\left(\frac{1}{2^{r}} - \frac{1}{3^{r}}\right).$$

By a similar argument, the result follows for the case when $-1 < \alpha < 0$ and $U_{[t(k)]} < U_{[s(k)]}$. Hence the proof. □

We have evaluated the coverage probabilities $\eta_{9(k)}(s,t,r;\alpha)$ under different combinations of (s,t), various values of α and r for $k = 1,2,3$. The values are presented in Table 4. It can be observed that coverage probabilities improve with the increase of α and k. Notice that when $-1 < \alpha < 0$, $\eta_{10(k)}(s,t,r;\alpha) = \eta_{9(k)}(s,t,r;-\alpha)$. Thus one can use Table 4 for evaluating $\eta_{10(k)}(s,t,r;\alpha)$.

Table 4: Values of $\eta_{9(k)}(s,t,r;\alpha)$ for some selected values of s,t,r and α

α	r	$k=1$			$k=2$			$k=3$		
		(s,t)			(s,t)			(s,t)		
		$(1,4)$	$(1,5)$	$(1,6)$	$(1,4)$	$(1,5)$	$(1,6)$	$(1,4)$	$(1,5)$	$(1,6)$
0.60	2	0.07292	0.07813	0.08073	0.07819	0.08916	0.09648	0.07227	0.08545	0.09534
0.75		0.09115	0.09766	0.10091	0.09774	0.11145	0.12060	0.09033	0.10681	0.11917
0.90		0.10938	0.11719	0.12109	0.11728	0.13374	0.14472	0.10840	0.12817	0.14301
0.95		0.11545	0.12370	0.12782	0.12380	0.14118	0.15276	0.11442	0.13529	0.15095
1.00		0.12153	0.13021	0.13455	0.13032	0.14861	0.16080	0.12045	0.14242	0.15890

6 Illustration Using Real Data

We consider a bivariate data set given in Platt et al. (1969) relating to 396 conifer (Pinus Palustris) trees. Chen et al. (2004) reproduced the data set as the first component X for a bivariate observation represents the diameter in centimeters of the conifer tree at breast height and the second component Y represents height in feet of the tree. Clearly X can be measured easily but it is somewhat difficult to measure Y. Also observations, such as girth (function of diameter) or height follows normal distribution. It is well known that logistic distribution is having more or less similar properties of a normal distribution(see, Malik, 1985) and hence it is known as an alternative model to normal distribution. We assume that (X,Y) follows Morgenstern type bivariate logistic distribution. By using the estimator of α given in Chacko and Thomas (2009), we take the estimator of α as $\hat{\alpha} = 1$. We have drawn a simple random sample of size 20 from the 396 conifer trees. Then the concomitants of order statistics arising from the sample are obtained as given below.

r	1	2	3	4	5	6	7	8	9	10
$Y_{[r:n]}$	5	3	3	5	8	14	15	20	22	17
r	11	12	13	14	15	16	17	18	19	20
$Y_{[r:n]}$	33	21	31	30	34	32	58	33	49	67

The concomitants of upper k-record values $(k = 1,2,3)$ extracted from the data set are obtained as given below. Based on the observed concomitants of order statistics and by using Table 1, we

n	1	2	3	4	5	6	7	8	9	10	11
$U_{[n(1)]}$	28	119	192	223	131	244	-	-	-	-	-
$U_{[n(2)]}$	28	26	43	119	192	222	223	208	131	-	-
$U_{[n(3)]}$	28	26	43	203	162	119	192	222	223	223	228

obtain the prediction intervals of future order statistics with prediction coefficient at least 30% are presented in the following.

n	r	(s,t)	$\left(Y_{[s:n]}, Y_{[t:n]}\right)$	$\eta_1\,(s,t,r;\alpha,m,n)$
15	8	$(4,19)$	$(5,49)$	0.33614
	8	$(1,19)$	$(5,49)$	0.40336
	12	$(1,19)$	$(5,49)$	0.30252

Based on the observed concomitant upper k-record values and by using Table 2, the prediction intervals of future order statistics with prediction coefficient at least 30% are presented in the following.

k	m	r	(s,t)	$\left(U_{[s(k)]}, U_{[t(k)]}\right)$	$\eta_{3(k)}\,(s,t,r;\alpha,m)$
2	20	10	$(1,8)$	$(28,208)$	0.3000
3	20	10	$(2,10)$	$(26,223)$	0.33033

On the basis of the coverage probability, it can be observed that the predictive intervals are improved with the increase of k.

7 Conclusion

In this paper, we developed distribution-free prediction intervals for the future order statistics and k-record values from an X sequence based on observed concomitants of order statistics and based on the observed concomitants of k-record values as Y sequence. These interval coverage probabilities obtained are accurate and independent of the parent distribution function. It can be observed that the coverage probabilities corresponds to order statistics improve with the increase of α and that of k-record values improve with the increase of α and k.

Acknowledgements

The authors would like to thank the editor and the anonymous reviewers for their helpful suggestions which substantially improved the earlier version of the manuscript.

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., Mir Mostafaee, 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.

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. (1992), *A First Course in Order Statistics*. USA: Classics in Applied Mathematics.

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

Bairamov, I. G. and Bekci, M. (1999), "Concomitant of order statistics in FGM type bivariate uniform distributions", *Istatistik, Journal of the Turkish Statistical Association*, 2(2), 135-144.

Balasubramanian, K. and Beg, M. I. (1997), "Concomitant of order statistics in Morgenstern type bivariate exponential distributions", *Journal of Applied Statistical Science*, 5, 233-245.

Balasubramanian, K. and Beg, M. I. (1998), "Concomitant of order statistics in Gumbel's bivariate exponential distribution", *Journal of Applied Statistical Science*, 5, 233-245.

Chacko, M. and Mary, S. (2013), "Concomitants of k-record values arising from Morgenstern family of distributions and their applications in parameter estimation", *Statistical Papers*, 54, 21-46.

Chacko, M., and Thomas, P. Y. (2009), "Estimation of parameters of Morgenstern type bivariate logistic distribution by ranked set sampling", *Journal of the Indian Society of Agricultural Statistics*, 63, 77-83.

Chacko, M. and Thomas, P. Y. (2011), "Estimation of parameter of Morgenstern type bivariate exponential distribution using concomitants of order statistics", *Statistical Methodology*, 8, 363-376.

Chen, Z., Bai, Z. and Sinha, B. K. (2004), "Lecture Notes in Statistics, Ranked Set Sampling", *Theory and Applications*, Springer, New York.

Cox, D. R. (1975), "Prediction intervals and empirical Bayes confidence intervals", *Journal of Applied Probability*, 12, 47-55.

David, H. A. (1973), "Concomitants of order statistics", *Bulletin of the International Statistical Institute*, 45, 295-300.

David, H. A. and Galambose, J. (1974), "The asymptotic theory of concomitants of order statistics", *Journal of Applied Probability*, 11, 762-770.

David, H. A., O'Connell, M. J., and Yang, S. S. (1977), "Distribution and expected value of the rank of a concomitant of an order statistic", *The Annals of Statistics*, 5, 216-223.

David, H. A. (1981), *Order Statistics*, New York: John Wiley and Sons.

David, H. A. and Nagaraja, H. N. (1998), "Concomitants of order statistics", in: N. Balakrishnan, C.R. Rao (Eds.), *Handbook of Statistics*, 16, Elsevier, Amsterdam.

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

Escobar, L. A. and Meeker, W. Q. (1999), "Statistical prediction based on censored life data", *Technometrics*, 41, 113-124.

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

Houchens, R. L. (1984), *Record value theory and inference*, Ph.D. Dissertation, University of California.

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

Kotz, S., Balakrishnan, N., and Johnson, N. L. (2000), *Distributions in statistics: continuous multivariate distribution*, 2nd edn. Wiley, New York.

Malik, H. J. (1985). *Logistic Distribution*, Encyclopedia of Statistical Sciences, 5, (eds S. Kotz and N.L. Johnson), John Wiley and Sons, New York.

Platt, W. J., Evans, G. M., and Rathbun, S. L. (1988), "The population dynamics of a long-lived Conifer (Pinus Palustris)", *The American Naturalist*, 131, 491-525.

Scaria, J. and Nair, N. U. (1999), "On concomitants of order statistics from Morgenstern family", *Biometrical Journal*, 4, 483-489.

Received: *24 March 2022*

Accepted: *7 September 2022*