
Prediction intervals for future order statistics from two independent sequences

M. M. Mohie El-Din¹, M. S. Kotb¹, W. S. Emam^{2, *}

¹Department of Mathematics, Faculty of Science, Al-Azhar University, Nasr City, Cairo, Egypt

²Department of Basic Science, Faculty of Engineering, British University in Egypt, Al-Shorouq City, Cairo, Egypt

Email address:

mmmmoheeldin@yahoo.com (M. M. M. El-Din), kottp5_mas@yahoo.com (M. S. Kotb), walid_emam42@yahoo.com (W. S. Emam)

To cite this article:

M. M. Mohie El-Din, M. S. Kotb, W. S. Emam. Prediction Intervals for Future Order Statistics from Two Independent Sequences. *American Journal of Theoretical and Applied Statistics*. Vol. 4, No. 1, 2015, pp. 33-40. doi: 10.11648/j.ajtas.20150401.16

Abstract: In this article, based on observed X -sequence of independent and identically distribution (iid) continuous random variables, we discuss the problem of predicting future order statistics from a Y -sequence of iid continuous random variables from the same distribution. Specifically, distribution-free prediction intervals (PIs) for an order statistic observation based on either progressive Type-II right censoring, or order data from the past X -sequence, as well as outer and inner PIs are derived based on order statistics observations. Such these intervals are exact and do not depend on the sampling distribution. Finally, a real life time data set that given to breakdown of an insulating fluid between electrodes is used to illustrate the proposed procedures.

Keywords: Distribution-Free Prediction Intervals, Order Statistics, Progressive Type-II Right Censoring, Coverage Probability

1. Introduction

In many statistical situations the observed data not appropriate with certain distribution, also there exist often many determining mistakes as determine the suitable distribution. Distribution-free prediction is useful procedure in prediction of the future observations without need to know the parent distribution, such these PIs that produced are exact and they do not depend on the sampling distribution. Prediction of future events on the basis of the past and present knowledge is a fundamental problem of statistics; there are different types of predictions of future observation such as one-sample prediction, two-sample prediction and multi sample prediction. We interested here on the two-sample prediction. In this prediction type, we use the first sequence to predict future observations from another independent sequence. The prediction in nature can be either parametric or nonparametric; we focus our attention on the nonparametric type that is known by distribution-free PIs. In recent years, distribution-free PIs are discussed by many researchers. Ahmadi, Raqab and Balakrishnan the most authors focusing on the distribution-free prediction problem. Ahmadi and Balakrishnan discussed distribution-free PIs for the quantiles of a distribution based on record ranges was introduced by Ref. [1], distribution-free PIs for the quantile intervals based on

current records in Ref. [2], distribution-free PIs for single order statistics based on records as well as PIs for a future records based on observed order statistics are obtained in Ref. [3], PIs for order statistics based on current record coverage in Ref. [4]. Recently, they established outer and inner PIs for order statistics based on current records in Ref. [5]. Ahmadi, MirMostafae and Balakrishnan in Ref. [6] established outer and inner PIs for records based on order statistics. Raqab in Ref. [7] derived PIs for a future current records based on current record. Raqab and Balakrishnan in Ref. [8] derived PIs for a future record, as well as outer and inner PIs for records interval based on records.

Sequel to the predictions for future order statistics, this paper discussed PIs for an order statistic based on two observed samples, progressively Type-II right censored sample and order statistics sample, as well as outer and inner PIs for order statistics interval based on order statistics are also discussed. Specifically, this paper organized as follows: Section 2 contains some preliminaries. In Section 3, we derive distribution-free PIs for single order statistic from a future Y -sequence of iid random variables, based on a progressively Type-II right censored sample and order statistics sample from the X -sequence. In Section 4, distribution-free outer and inner PIs for future order statistics intervals from Y -sequence based on also order statistics observations from the X -sequence are

derived. In Section 5, numerical results to illustrate all the results derived here. Finally, a conclusion of this study is given in Section 6.

2. Preliminaries

Let X_1, X_2, \dots, X_n and Y_1, Y_2, \dots, Y_m denote two sequences of the lifetimes of reliability experiment units that placed on a life-test, we shall assume that these variables of the two samples are iid from an absolutely continuous population with cumulative distribution function (cdf) F and probability density function (pdf) f . Suppose based on the past knowledge of X-observations, we wish to construct a two-sided distribution-free PIs for the r^{th} future order lifetime $Y_{r:m}, 1 \leq r \leq m$ and also outer and inner PIs for a future order interval $(Y_{r:m}, Y_{s:m}), 1 \leq r \leq s \leq m$ from the future usual order statistics $Y_{1:m} \leq Y_{2:m} \leq \dots \leq Y_{m:m}$. Order statistics play a key role in studying the lifetimes of such systems, and it was raised in many practical situations including characterization of probability distributions and goodness-of-fit tests, analysis of censored samples, reliability analysis; for more details concerning order statistics see Arnold *et al.* [9] and David and Nagaraja [10]. The marginal pdf and the survival function of $Y_{r:m}$ and the joint pdf of any two order statistics, say $Y_{r:m} < Y_{s:m}$ such that $1 < r < s < m$, are easily expressed in terms of $F(\cdot)$ and $f(\cdot)$ as

$$f_{Y_{r:m}}(y) = r \binom{m}{r} [F(y)]^{r-1} [\bar{F}(y)]^{m-r} f(y), \quad (1)$$

$$\bar{F}_{Y_{r:m}}(y) = \sum_{\ell=0}^{r-1} \binom{m}{\ell} [F(y)]^{\ell} [\bar{F}(y)]^{m-\ell}, \quad (2)$$

and

$$f_{Y_{r:m}, Y_{s:m}}(y_r, y_s) = m! \frac{[F(y_r)]^{r-1}}{(r-1)!} f(y_r) \frac{[F(y_s) - F(y_r)]^{s-r-1}}{(s-r-1)!} f(y_s) \frac{[\bar{F}(y_s)]^{m-s}}{(m-s)!}, \quad (3)$$

respectively, where the survival function $\bar{F}(0) = 1 - F(0)$, $1 < r < s < m$, we need to referred that all the prediction coefficients were established here depend also in the observed X-sequence size n and the future Y-sequence size m .

3. PIs for Individual Order Statistics

In this section, we obtain two-sided distribution-free PIs for a future r^{th} order statistic $Y_{r:m}, 1 \leq i \leq m$ based on the endpoints of observed progressively Type-II right censored and order statistics, consecutively.

3.1. Based on progressive Type-II Right Censored

Suppose X_1, X_2, \dots, X_n be an iid random variables (r.v.'s) from a continuous cdf $F(x)$, and $X_{1:\mu:n} < \dots < X_{\mu:\mu:n}$ describe the progressively Type-II right-censored order

statistics failure times of size μ that observed from these r.v.'s with (R_1, \dots, R_{μ}) as progressive censoring scheme. For more details concerning about a progressively Type-II right censored order statistics sample, see Balakrishnan [11], here the discussion of distribution-free PIs for an order statistic from a future Y-sequence of iid r.v.'s from a continuous cdf $F(x)$, based on a past progressive Type-II right-censored order statistics. Balakrishnan *et al.* [12] derived an expression for the pdf and the survival function of $X_{i:\mu:n}$ ($1 \leq i \leq \mu \leq n$), as

$$f_{X_{i:\mu:n}}(x_i) = c_i' \sum_{\ell=0}^{i-1} c_{\ell, i-1}(R_{i-1}) f(x_i) [\bar{F}(x_i)]^{R_{\ell, i}^{\nu} - 1}, \quad (4)$$

and

$$\bar{F}_{X_{i:\mu:n}}(x_i) = c_i' \sum_{\ell=0}^{i-1} \frac{c_{\ell, i-1}(R_{i-1})}{R_{\ell, i}^{\nu}} [\bar{F}(x_i)]^{R_{\ell, i}^{\nu}}, \quad (5)$$

respectively, where $-\infty < x_i < \infty$,

$$c_i' = n(n - R_1 - 1) \dots (n - R_1 - \dots - R_{i-1} - i + 1),$$

$$R_{i-1} = (R_1 + 1, \dots, R_{i-1} + 1), \quad R_{\ell, i}^{\nu} = (R_i^* + 1) + \sum_{j=i-\ell}^{i-1} (R_j + 1),$$

$$R_i^* = n - i - R_1 - \dots - R_{i-1}, \text{ and}$$

$$c_{\ell, i}(R_i) = \frac{(-1)^{\ell}}{\left(\prod_{j=1}^{\ell} \sum_{k=i-\ell+1}^{i-\ell+j} (R_k + 1) \right) \left(\prod_{j=1}^{i-\ell} \sum_{k=j}^{i-\ell} (R_k + 1) \right)}.$$

Theorem 1. Let $X_{j:\mu:n}$ ($1 \leq j \leq \mu \leq n$) be progressive Type-II right-censored order statistics based on continuous cdf $F(x)$. Moreover, let $Y_{1:m} \leq Y_{2:m} \leq \dots \leq Y_{m:m}$ be the order statistics from a future random sample of size m from the same cdf $F(y)$. Then $(X_{i:\mu:n}, X_{j:\mu:n}), 1 \leq i < j \leq \mu \leq n$, is a two-sided PI for $Y_{r:m}, 1 \leq r \leq m$, whose coverage probability is free of the parent distribution F and is given by

$$\pi_1(i, j, r) = \frac{m! c_j'}{(m-r)!} \sum_{\ell=0}^{j-1} \frac{c_{\ell, j-1}(R_{j-1})}{R_{j, \ell}^{\nu} (R_{j, \ell}^{\nu} + m) \dots (R_{j, \ell}^{\nu} + m - r + 1)} - \frac{m! c_i'}{(m-r)!} \sum_{\ell=0}^{i-1} \frac{c_{\ell, i-1}(R_{i-1})}{R_{i, \ell}^{\nu} (R_{i, \ell}^{\nu} + m) \dots (R_{i, \ell}^{\nu} + m - r + 1)}. \quad (6)$$

Proof: Under the assumption that $X_{j:\mu:n}, 1 \leq j \leq \mu \leq n$ are continuous r.v.'s, we can write

$$P(X_{i:\mu:n} \leq y \leq X_{j:\mu:n}) = P(X_{i:\mu:n} \leq y) - P(X_{j:\mu:n} \leq y). \quad (7)$$

From (5), we readily obtain

$$P(X_{i:\mu:n} \leq y \leq X_{j:\mu:n}) = c_j' \sum_{\ell=0}^{j-1} \frac{c_{\ell, j-1}(R_{j-1})}{R_{j, \ell}^{\nu}} [\bar{F}(x_j)]^{R_{j, \ell}^{\nu}} - c_i' \sum_{\ell=0}^{i-1} \frac{c_{\ell, i-1}(R_{i-1})}{R_{i, \ell}^{\nu}} [\bar{F}(x_i)]^{R_{i, \ell}^{\nu}}. \quad (8)$$

Using the conditioning argument, we then have

$$P(X_{i:\mu:n} \leq Y_{r:m} \leq X_{j:\mu:n}) = \int_{-\infty}^{\infty} P(X_{i:\mu:n} \leq Y_{r:m} \leq X_{j:\mu:n} | Y_{r:m} = y) dF_{Y_{r:m}}(y) = \int_{-\infty}^{\infty} P(X_{i:\mu:n} \leq y \leq X_{j:\mu:n}) f_{r,m}(y) dy \tag{9}$$

Upon using (1) and (8) in (9), we obtain

$$P(X_{i:\mu:n} \leq Y_{r:m} \leq X_{j:\mu:n}) = r \binom{m}{r} \left(c_j \sum_{\ell=0}^{j-1} \frac{c_{\ell,j-1} \binom{R_{j-1}}{R_{j,\ell}} I_{\ell}(j) + c_i \sum_{\ell=0}^{i-1} \frac{c_{\ell,i-1} \binom{R_{i-1}}{R_{i,\ell}} I_{\ell}(i) \right) \tag{10}$$

where

$$I_{\ell}(j) = \int_{-\infty}^{\infty} (F(y))^{r-1} (1-F(y))^{R_{j,\ell}+u-r} f(y) dy = B(r, R_{j,\ell} + u - r + 1) \tag{11}$$

By expanding previous beta constant, easily we obtain the required result. Under the assumptions of theorem 1, we note that $X_{i:\mu:n}$ is a lower prediction bound and $X_{j:\mu:n}$ is an upper bound for $Y_{r:m}$, $1 \leq r \leq m$. Note that, the choice of i and j is not unique. So, for a given confidence level π_0 and specified r and m , we would like to construct a PI as short as possible among all PIs with the same level. Table 1 and 2 presents values of $\pi_1(i, j; r)$ for $r = 5, 10, 15$ and $m = 50$ for some choices of i, j, μ, n and the progressive censoring scheme. Thus, $(X_{i:\mu:n}, X_{j:\mu:n})$ be a $(\pi_1)100\%$ prediction interval for $Y_{r:m}$, the r^{th} order statistic from a future sample of size m , with $\pi_1(i, j; r)$ given by (6) which does not depend on the parent distribution F .

Table 1. Values of $\pi_1(i, j; r)$ for some choices of i, j, r , when $n = 50, \mu = 40$, progressive censoring scheme $(0, 2, 0, 5, 0, 3, 0, \dots, 0)$ and $m = 50$.

r	i	j					
		10	15	20	25	30	40
5	1	0.971	0.972	0.972	0.972	0.972	0.972
	2	0.857	0.897	0.898	0.898	0.898	0.898
	3	0.736	0.775	0.776	0.776	0.776	0.776
	4	0.159	0.313	0.322	0.323	0.323	0.323
10	1	0.395	0.874	0.972	0.993	0.994	0.994
	2	0.368	0.847	0.945	0.966	0.967	0.967
	3	0.298	0.777	0.876	0.897	0.898	0.898
15	4	0.184	0.663	0.762	0.783	0.784	0.784
	1	0.111	0.593	0.863	0.989	0.999	1.000
	2	0.109	0.590	0.860	0.987	0.996	0.997
	3	0.099	0.580	0.850	0.976	0.986	0.987
4	0.073	0.554	0.824	0.950	0.960	0.961	

Remark: By specially choosing for $r = 1$ and m , in equation (6), we can formed the coverage probability of the minimum and the maximum of the future order sample, based on progressive Type-II right-censored order statistics, respectively as respectively as

$$\pi_1(i, j; 1) = mc_j \sum_{\ell=0}^{j-1} \frac{c_{\ell,j-1} \binom{R_{j-1}}{R_{j,\ell}}}{R_{j,\ell} (R_{j,\ell} + m)} - mc_i \sum_{\ell=0}^{i-1} \frac{c_{\ell,i-1} \binom{R_{i-1}}{R_{i,\ell}}}{R_{i,\ell} (R_{i,\ell} + m)},$$

$$\pi_1(i, j; m) = mlc_j \sum_{\ell=0}^{j-1} \frac{c_{\ell,j-1} \binom{R_{j-1}}{R_{j,\ell}}}{R_{j,\ell} (R_{j,\ell} + m) \dots (R_{j,\ell} + 1)} - mlc_i \sum_{\ell=0}^{i-1} \frac{c_{\ell,i-1} \binom{R_{i-1}}{R_{i,\ell}}}{R_{i,\ell} (R_{i,\ell} + m) \dots (R_{i,\ell} + 1)}$$

3.2. Based on Order Statistics

Let X_1, X_2, \dots, X_n lifetimes units, these lifetimes are iid from a continuous population with cdf $F(x)$ and pdf $f(x)$. $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ denote the corresponding ordered lifetimes, the pdf and the survival function of j^{th} order statistics $X_{j:n}$, $1 \leq j \leq n$ are given by (1) and (2) respectively. Here we derived two-sided distribution-free PIs for a future r^{th} order statistic $Y_{r:m}$, $1 \leq i \leq m$ based on the endpoints from these X-ordered lifetimes.

Table 2. Values of $\pi_1(i, j; r)$ for some choices of i, j, r , when $n = 30, \mu = 25$, progressive censoring scheme $(0, 2, 0, 3, 0, \dots, 0)$ and $m = 50$.

r	i	j					
		5	8	10	13	15	20
5	1	0.748	0.903	0.911	0.912	0.912	0.912
	2	0.572	0.726	0.735	0.736	0.736	0.736
	3	0.351	0.506	0.514	0.515	0.515	0.515
	4	0.590	0.629	0.630	0.630	0.630	0.630
10	1	0.637	0.955	0.998	0.999	0.999	0.999
	2	0.634	0.651	0.994	0.996	0.996	0.996
	3	0.623	0.940	0.983	0.985	0.985	0.985
15	4	0.599	0.916	0.959	0.961	0.961	0.961
	1	0.221	0.720	0.965	0.999	1.000	1.000
	2	0.221	0.720	0.965	0.999	1.000	1.000
	3	0.220	0.720	0.964	0.998	0.999	0.999
4	0.219	0.718	0.963	0.997	0.998	0.998	

Theorem 2. Let $Y_{i:m}$, $(1 \leq i \leq m)$ be i^{th} order statistic from a future random sample of size m with cdf $F(y)$. If $X_{j:n}$ is the j^{th} order statistics from another observed random sample of size n with the same cdf $F(x)$, then $(X_{i:n}, X_{j:n})$, $1 \leq i < j \leq n$, is a two-sided PI for $Y_{r:m}$, $1 \leq r \leq m$, whose coverage probability is free of F and is given by

$$\pi_2(i, j; r) = r \binom{m}{r} \sum_{\ell=1}^{j-1} \binom{n}{\ell} \left\{ (\ell+r) \binom{n+m}{\ell+r} \right\}^{-1} \tag{12}$$

Proof: Similarly (7) and from (2) when X is continuous, we have

$$P(X_{i:n} \leq y \leq X_{j:n}) = \sum_{\ell=1}^{j-1} \binom{n}{\ell} (F(y))^\ell (1-F(y))^{n-\ell} \tag{13}$$

We use the same method as in the proof of Theorem 1; by using the conditioning argument on $Y_{r:m} = y$ and the fact

$$P(X_{i:n} \leq Y_{r:m} \leq X_{j:n}) = \int_{-\infty}^{\infty} P(X_{i:n} \leq y \leq X_{j:n}) dF_{r:m}(y).$$

Substituting (1) and (13) in (14), we obtain

$$\begin{aligned} &P(X_{i:n} \leq Y_{r:m} \leq X_{j:n}) \\ &= r \binom{m}{r} \sum_{\ell=i}^{j-1} \binom{n}{\ell} \int_{-\infty}^{\infty} (F(y))^{r+\ell-1} (1-F(y))^{n+m-r-\ell} f(y) dy \\ &= r \binom{m}{r} \sum_{\ell=i}^{j-1} \binom{n}{\ell} \int_0^1 u^{r+\ell-1} (1-u)^{n+m-r-\ell} du \\ &= r \binom{m}{r} \sum_{\ell=i}^{j-1} \binom{n}{\ell} B(r+\ell, n+m-r-\ell+1). \end{aligned}$$

As the previous section, by expanding beta constant, we can obtain the required result. Table 3 presents values of

Table 3. The values of $\pi_2(i, j; r)$ for some selected choices of n, m, i, j and r .

(n, m)	r	i	j						n
			r+1	r+2	r+3	r+4	r+5	r+10	
(30,20)	5	r-4	0.348	0.460	0.568	0.666	0.750	0.962	0.993
		r-3	0.323	0.435	0.544	0.642	0.726	0.937	0.968
		r-2	0.275	0.387	0.495	0.594	0.678	0.889	0.920
	10	r-4	0.174	0.246	0.328	0.417	0.509	0.876	0.986
		r-3	0.160	0.232	0.314	0.403	0.495	0.862	0.972
		r-2	0.138	0.209	0.291	0.380	0.472	0.840	0.950
	15	r-4	0.065	0.101	0.149	0.211	0.287	0.775	0.993
		r-3	0.061	0.097	0.145	0.207	0.283	0.771	0.989
		r-2	0.054	0.090	0.138	0.200	0.276	0.764	0.982
(50,40)	20	r-4	0.188	0.250	0.318	0.391	0.466	0.791	0.971
		r-3	0.166	0.228	0.297	0.370	0.445	0.769	0.935
		r-2	0.137	0.199	0.268	0.340	0.415	0.740	0.906
	25	r-4	0.138	0.189	0.248	0.315	0.387	0.751	0.972
		r-3	0.123	0.174	0.234	0.300	0.372	0.737	0.958
		r-2	0.103	0.154	0.214	0.280	0.353	0.717	0.937
	30	r-4	0.087	0.126	0.173	0.231	0.299	0.708	0.987
		r-3	0.080	0.118	0.166	0.223	0.291	0.700	0.979
		r-2	0.068	0.106	0.154	0.212	0.279	0.689	0.967
(100,90)	40	r-4	0.200	0.251	0.305	0.360	0.415	0.662	0.662
		r-3	0.168	0.220	0.274	0.329	0.383	0.631	0.850
		r-2	0.133	0.184	0.238	0.293	0.348	0.595	0.815
	60	r-4	0.144	0.188	0.235	0.286	0.341	0.624	0.933
		r-3	0.125	0.168	0.215	0.266	0.321	0.604	0.913
		r-2	0.100	0.143	0.191	0.242	0.297	0.580	0.889
	80	r-4	0.058	0.084	0.118	0.160	0.211	0.591	0.989
		r-3	0.053	0.079	0.112	0.155	0.206	0.586	0.984
		r-2	0.045	0.071	0.105	0.150	0.198	0.578	0.967

4. PIs for Order Interval

In this section, we obtain outer and inner two-sided distribution-free PIs for a future order interval $(Y_{r:m}, Y_{s:m})$ based on also order statistics.

4.1. Outer PIs Based on Order Statistics

Suppose $(Y_{r:m}, Y_{s:m})$ is a future order interval from the Y-sequence and $X_{j:n}$, $1 \leq j \leq n$ be the j^{th} order statistics from X-sequence, we are interested here in obtaining $100(1 - \alpha)\%$ PIs for it of the form $(X_{i:n}, X_{j:n})$ such that

$\pi_2(i, j; r)$ for some choices of n, m, i, j and r . Thus, $(X_{i:n}, X_{j:n})$ is a $(\pi_2)100\%$ PI for the future r^{th} order statistic $Y_{r:m}$, $1 \leq i \leq m$, with $\pi_2(i, j; r)$ given by (12) which does not depend on the parent distribution F .

Also, we can form the coverage probability of the minimum and the maximum of the future order sample, based on ordered lifetime observations by sitting $r = 1, m$, respectively in (12).

$p(X_{i:n} \leq Y_{r:m} \leq Y_{s:m} \leq X_{j:n}) \geq 1 - \alpha$. By using the binomial expansion we can reformed the joint pdf of any two order statistics $X_{i:n} < X_{j:n}$, $1 < i < j < n$ that given by (3) as

$$f_{X_{i:n}, X_{j:n}}(x_i, x_j) = \frac{n!}{(i-1)!} \sum_{k=0}^{j-1-i} \sum_{z=0}^{n-j-k} a_{k,z} \left(\frac{F(x_i)}{F(x_j)} \right)^{i+k} F(x_j)^{j+z} \frac{f(x_i)f(x_j)}{F(x_i)F(x_j)}, \quad (14)$$

where

$$a_{k,z} = \frac{(-1)^{k+z}}{(j-i-k-1)!k!(n-j-z)!z!}. \quad (15)$$

Theorem 3. Suppose the conditions of Theorem 2 hold, Then, $(X_{i:n}, X_{j:n})$ is an outer PIs for the future order interval $(Y_{r:m}, Y_{s:m})$ from the Y-sequence, with the corresponding prediction coefficient, being free of F , given by

$$\pi_3(i, j; r, s) = \frac{n!m!}{(i-1)!(r-1)!} \sum_{k=0}^{j-i-1} \sum_{z=0}^{n-j} a_{k,z} \sum_{q=0}^{s-r-1} \sum_{w=0}^{m-s} \frac{b_{q,w}}{i+k+r+q} \times \frac{1}{(i+k+s+w)(j+z+s+w)}, \quad (16)$$

where $a_{k,z}$ is given by Equation (15) and

By using (14) and (19) we also obtain

$$p(X_{i:n} \leq y_r \leq y_s \leq X_{j:n}) = \frac{n!}{(i-1)!} \sum_{k=0}^{j-i-1} \sum_{z=0}^{n-j} a_{k,z} \int_{F(y_s)}^1 v^{j-i-k+z-1} dv \int_0^{F(y_r)} u^{i+k-1} du = \frac{n!}{(i-1)!} \sum_{k=0}^{j-i-1} \sum_{z=0}^{n-j} a_{k,z} \frac{(1-F(y_s))^{j-i-k+z} F(y_r)^{i+k}}{(j-i-k+z)(i+k)}. \quad (19)$$

$$p(X_{i:n} \leq Y_{r:m} \leq Y_{s:m} \leq X_{j:n}) = \int_{-\infty}^{\infty} \int_{y_r}^{\infty} p(X_{i:n} \leq y_r \leq y_s \leq X_{j:n}) f_{Y_{r:m}, Y_{s:m}}(y_r, y_s) dy_s dy_r = \int_{-\infty}^{\infty} \int_{y_r}^{\infty} p(X_{i:n} \leq y_r \leq y_s \leq X_{j:n}) f_{Y_{r:m}, Y_{s:m}}(y_r, y_s) dy_s dy_r = \frac{n!m!}{(i-1)!(r-1)!} \sum_{k=0}^{j-i-1} \sum_{z=0}^{n-j} \frac{a_{k,z}}{(j-i-k+z)(i+k)} \sum_{q=0}^{s-r-1} \sum_{w=0}^{m-s} b_{q,w} \int_0^1 \int_v^1 (1-u)^{j-i-k+z} u^{s-r-q+w-1} v^{i+k+r+q-1} dudv = \frac{n!m!}{(i-1)!(r-1)!} \sum_{k=0}^{j-i-1} \sum_{z=0}^{n-j} a_{k,z} \sum_{q=0}^{s-r-1} \sum_{w=0}^{m-s} \frac{b_{q,w}}{i+k+r+q} \frac{1}{(i+k+s+w)(j+z+s+w)}. \quad (20)$$

Table 4. Values of $\pi_3(i, j; r, s)$ for some choices of i, j, r and s , such $n = 50$ and $m = 30$.

r	s	i	j					
			20	25	30	35	40	45
3	9	1	0.786	0.917	0.947	0.950	0.951	0.951
		2	0.695	0.822	0.851	0.854	0.854	0.854
		3	0.580	0.699	0.727	0.730	0.730	0.730
		4	0.460	0.568	0.595	0.598	0.598	0.598
5	9	1	0.828	0.961	0.990	0.993	0.994	0.994
		2	0.808	0.941	0.971	0.974	0.974	0.974
		3	0.770	0.902	0.931	0.935	0.935	0.935
		4	0.711	0.842	0.871	0.875	0.875	0.875
7	9	1	0.833	0.966	0.996	0.999	0.999	0.999
		2	0.830	0.963	0.993	0.996	0.996	0.996
		3	0.821	0.954	0.984	0.988	0.988	0.988
		4	0.804	0.938	0.967	0.971	0.971	0.971
3	11	1	0.596	0.842	0.932	0.949	0.951	0.951
		2	0.519	0.750	0.837	0.953	0.854	0.854
		3	0.422	0.632	0.713	0.729	0.730	0.730
		4	0.324	0.508	0.582	0.596	0.598	0.598
5	11	1	0.632	0.884	0.976	0.993	0.994	0.994
		2	0.614	0.864	0.956	0.973	0.974	0.974
		3	0.579	0.826	0.917	0.934	0.935	0.935
		4	0.526	0.767	0.857	0.873	0.875	0.875
7	11	1	0.637	0.889	0.981	0.998	0.999	0.999
		2	0.634	0.886	0.978	0.995	0.996	0.996
		3	0.626	0.877	0.970	0.985	0.988	0.988
		4	0.610	0.860	0.953	0.970	0.971	0.971

$$b_{q,w} = \frac{(-1)^{q+w}}{(s-r-q-1)!k!(m-s-w)!w!}. \quad (17)$$

Proof: Under the assumption that the j^{th} observed order statistics $X_{j:n}, 1 \leq j \leq n$ are continuous r.v.'s, we can write

$$p(X_{i:n} \leq y_r \leq y_s \leq X_{j:n}) = \int_{y_s}^{\infty} \int_{-\infty}^{y_r} f_{X_{i:n}, X_{j:n}}(x_i, x_j) dx_i dx_j, \quad (18)$$

By the earlier reformed Equation (14), we can easily reduced Equation (18) to

We realize that $\pi_3(i, j; r, s)$ does not depend on F . Some values of $\pi_3(i, j; r)$ are presents in table 4 for $n = 50$ and $m = 30$ for some choices of i, j, r and s . Where, $(X_{i:n}, X_{j:n})$ be a $(\pi_3)100\%$ outer prediction interval for $(Y_{r:m}, Y_{s:m})$, such $\pi_3(i, j; r, s)$ given by (16).

4.2. Inner PIs Based on Order Statistics

Suppose we are interested in obtaining $100(1 - \alpha)\%$ inner two-sided distribution-free PIs for a future order interval $(Y_{r:m}, Y_{s:m})$ based on order statistics of the form $(X_{i:n}, X_{j:n})$ such that $p(Y_{r:m} \leq X_{i:n} \leq X_{j:n} \leq Y_{s:m}) \geq 1 - \alpha$. Here, we describe how such distribution-free inner PIs can be constructed.

Theorem 4. Suppose the conditions of Theorem2 hold, then, $(X_{i:n}, X_{j:n})$ is an inner PIs for the order interval $(Y_{r:m}, Y_{s:m})$ from the Y-sequence, with the corresponding prediction coefficient does not depend on the sampling distribution, given by

$$\pi_4(i, j; r, s) = \frac{n!m!}{(i-1)!(r-1)!} \sum_{k=0}^{j-i-1} \sum_{z=0}^{n-j} a_{k,z} \sum_{q=0}^{s-r-1} \sum_{w=0}^{m-s} b_{q,w} \times \frac{1}{(q+r)(i+k+q+r)(j+q+r+z)(j+s+w+z)}, \tag{21}$$

i.v.'s, we can write

$$p(y_r \leq X_{i:n} \leq X_{j:n} \leq y_s) = \int_{y_r}^{y_s} \int_{y_r}^{x_j} f_{X_{i:n}, X_{j:n}}(x_i, x_j) dx_i dx_j. \tag{22}$$

Proceeding similarly, from (14) in (22), we obtain

where $a_{k,z}$ and $b_{q,w}$ are given respectively, by Equations (15) and (17).

Proof: For a given $y_r \leq y_s$, under the assumption that the j^{th} observed order statistics $X_{j:n}$, $1 \leq j \leq n$ are continuous

$$\begin{aligned} p(y_r \leq X_{i:n} \leq X_{j:n} \leq y_s) &= \frac{n!}{(i-1)!} \sum_{k=0}^{j-i-1} \sum_{z=0}^{n-j} a_{k,z} \int_{y_r}^{y_s} \int_{y_r}^{x_j} [F(x_i)]^{i+k-1} [F(x_j)]^{j-i-k+z-1} f(x_i) f(x_j) dx_i dx_j \\ &= \frac{n!}{(i-1)!} \sum_{k=0}^{j-i-1} \sum_{z=0}^{n-j} a_{k,z} \int_{F(y_r)}^{F(y_s)} \int_{F(y_r)}^v u^{i+k-1} v^{j-i-k+z-1} dudv \\ &= \frac{n!}{(i-1)!} \sum_{k=0}^{j-i-1} \sum_{z=0}^{n-j} \frac{a_{k,z}}{(i+k)} \left(\frac{F(y_s)^{j+z} - F(y_r)^{j+z}}{j+z} - \frac{F(y_r)^{i+k} (F(y_s)^{j-i-k+z} - F(y_r)^{j-i-k+z})}{j-i-k+z} \right). \end{aligned} \tag{23}$$

And by using (23) and (14), it is now known to verify the following conditioning argument

$$\begin{aligned} p(Y_{r:m} \leq X_{i:n} \leq X_{j:n} \leq Y_{s:m}) &= \int_{-\infty}^{\infty} \int_{y_r}^{\infty} p(y_r \leq X_{i:n} \leq X_{j:n} \leq y_s) f_{Y_{r:m}, Y_{s:m}}(y_r, y_s) dy_s dy_r \\ &= \frac{n!m!}{(i-1)!(r-1)!} \sum_{k=0}^{j-i-1} \sum_{z=0}^{n-j} \frac{a_{k,z}}{(i+k)} \sum_{q=0}^{s-r-1} \sum_{w=0}^{m-s} b_{q,w} \int_0^1 \int_r^1 \left(\frac{p^{j+z} - t^{j+z}}{j+z} - \frac{t^{i+k} (p^{j-i-k+z} - t^{j-i-k+z})}{j-i-k+z} \right) t^{r+q-1} p^{s-r-q+w-1} dp dt \\ &= \frac{n!m!}{(i-1)!(r-1)!} \sum_{k=0}^{j-i-1} \sum_{z=0}^{n-j} a_{k,z} \sum_{q=0}^{s-r-1} \sum_{w=0}^{m-s} b_{q,w} \frac{1}{(q+r)(i+k+q+r)(j+q+r+z)(j+s+w+z)}. \end{aligned} \tag{24}$$

Some values of $\pi_4(i, j; r)$ which given by (21) are presents in table 5 for some $n, m=30, 40, 50, 40$, consecutively, and for some choices of i, j, r and s . Thus, $(X_{i:n}, X_{j:n})$ is a $(\pi_4)100\%$ inner prediction interval for $(Y_{r:m}, Y_{s:m})$.

Table 5. Values of $\pi_4(i, j; r, s)$ for some choices of i, j, r and s .

r	s	i	j n=30, m=40				j n=50, m=40			
			r+8	r+10	r+12	r+15	r+8	r+10	r+12	r+15
1	15	r+1	0.599	0.416	0.240	0.069	0.677	0.647	0.591	0.458
		r+5	0.752	0.537	0.318	0.097	0.956	0.920	0.851	0.677
		r+7	0.756	0.541	0.321	0.098	0.974	0.639	0.869	0.694
1	20	r+1	0.781	0.707	0.574	0.312	0.694	0.692	0.686	0.658
		r+5	0.954	0.871	0.717	0.401	0.974	0.972	0.965	0.932
		r+7	0.958	0.875	0.721	0.404	0.992	0.990	0.983	0.950
1	30	r+1	0.820	0.819	0.817	0.794	0.694	0.694	0.694	0.694
		r+5	0.995	0.995	0.992	0.967	0.974	0.974	0.974	0.974
		r+7	0.999	0.999	0.996	0.971	0.993	0.993	0.993	0.993
3	25	r+1	0.775	0.740	0.661	0.443	0.550	0.550	0.549	0.544
		r+5	0.972	0.935	0.846	0.586	0.907	0.907	0.906	0.901
		r+7	0.983	0.946	0.857	0.596	0.965	0.965	0.964	0.958
4	30	r+1	0.791	0.786	0.766	0.666	0.512	0.512	0.512	0.511
		r+5	0.983	0.977	0.956	0.844	0.874	0.874	0.874	0.874
		r+7	0.995	0.990	0.969	0.856	0.947	0.947	0.947	0.947
5	40	r+1	0.799	0.799	0.799	0.799	0.481	0.481	0.481	0.481
		r+5	0.983	0.983	0.983	0.983	0.843	0.843	0.843	0.843
		r+7	0.996	0.996	0.996	0.996	0.928	0.928	0.928	0.928

5. Illustrative Example (Real Life Data)

In this section, we consider the real life data set which given

in Nelson [20] to illustrate the methods proposed in the previous sections. These data which was also used in Lawless ([19], p. 185), concerning the data on time to breakdown of an insulating fluid between electrodes at a voltage of 34 kV

(minutes). The 19 times to breakdown are contained in the sample (*)

0.96 4.15 0.19 0.78 8.01 31.75 7.35 6.50 8.27 33.91
32.52 3.16 4.85 2.78 4.67 1.31 12.06 36.71 72.89

Also, from the data (*), we observe the following order statistics values:

0.19 0.78 0.96 1.31 2.78 3.16 4.15 4.67 4.85 6.50
7.35 8.01 8.27 12.06 31.75 32.52 33.91 36.71 72.89

Moreover, let $\mu : n = 14 : 19$ and the progressive censoring scheme $(0,2,0,3,0, \dots, 0)$. The observed progressive Type-II right-censored order statistics from the data (*) will be:

0.19 0.96 1.31 2.78 3.16 4.15 4.67
6.50 7.35 8.01 8.27 32.52 33.91 36.71

Table 6 presents values of π_1 and π_2 based on the above the real life data for some choices of i, j and r . That given the distribution-free PIs for future r^{th} order statistics $Y_{r:10}, 1 \leq i \leq 10$, from order sample of size $m = 10$, based on previous observed progressive Type-II right-censored and

order statistics, respectively.

6. Conclusion

In this article, we have derived distribution-free PIs for future order statistics based on progressively Type-II right censored sample and for future order statistics based on order statistics. Also, outer and inner PIs are derived based on order statistics observations. In general, the prediction coefficients values increase as the length of the PIs increase, else the prediction coefficients values of the inner prediction case it increase as the PIs length decrease. It may be noted from the results in tables 1 and 2, the prediction coefficients for a future r^{th} order statistic $Y_{r:m}, 1 \leq i \leq m$ based on the endpoints of observed progressively Type-II right censored increase as m increase, even if the censored observations size $n - m$ increase. For example, at the same size of the future order $m = 50$, we find the results of table 1 are greater than the corresponding results in table 2 although the censored observations in it increased.

Table 6. PIs for future order statistics based on above observed progressive Type-II right-censored and order statistics.

r	(i, j)	$(X_{i:14:19}, X_{j:14:19})$	π_1	$(X_{i:19}, X_{j:19})$	π_2
	(1, 8)	(0.19, 6.50)	0.6488	(0.19, 4.67)	0.6376
	(2, 9)	(0.96, 7.35)	0.7174	(0.78, 4.85)	0.6824
	(2, 8)	(0.96, 6.50)	0.7867	(0.78, 4.67)	0.6746
	(1, 9)	(0.19, 7.35)	0.8526	(0.19, 4.85)	0.6819
	(1, 10)	(0.19, 8.01)	0.8342	(0.19, 6.50)	0.5979
	(4, 11)	(2.78, 8.27)	0.7875	(1.31, 7.35)	0.4847
	(4, 13)	(2.78, (33.91))	0.9029	(1.31, 8.01)	0.5162
	(5, 14)	(3.16, 36.71)	0.9242	(2.78, 12.06)	0.4159
	(7, 14)	(4.67, (36.71))	0.8066	(4.15, 12.06)	0.2025
	(7, 14)	(4.67, (36.71))	0.5618	(4.15, 12.06)	0.0567

References

[1] J. Ahmadi, N. Balakrishnan, Confidence intervals for quantiles in terms of record range, Stat. Prob. Letters 68 (2004) 395 – 405.

[2] J. Ahmadi, N. Balakrishnan, Distribution-free confidence intervals for quantile intervals based on current records, Stat. Prob. Letters 75 (2005) 190 – 202.

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

[4] J. Ahmadi, N. Balakrishnan, Distribution-free prediction intervals for order statistics based on record coverage, J. Korean Stat. Society 40 (2011) 181 – 192.

[5] J. Ahmadi, N. Balakrishnan, Outer and inner prediction intervals for order statistics based on current records, Springer 53 (2012) 789– 802.

[6] J. Ahmadi, S.M.T.K. MirMostafae, N. Balakrishnan, Nonparametric prediction intervals for future record intervals based on order statistics, Stat. Prob. Letters 80 (2010) 1663 – 1672.

[7] M.Z. Raqab, Distribution-free prediction intervals for the future current record statistics, Springer 50 (2009) 429– 439.

[8] M. Z. Raqab, N. Balakrishnan, Prediction intervals for future records, Stat. Prob. Letters, 78 (2008) 1955 – 1963.

[9] B.C. Arnold, N. Balakrishnan, H.N. Nagaraja, A First Course in Order Statistics, John Wiley - Sons, New York (1992).

[10] H.A. David, H.N. Nagaraja, Order Statistics, third ed.. John Wiley - Sons, Hoboken, New Jersey (2008).

[11] N. Balakrishnan, Progressive censoring methodology: an appraisal, Springer, 16 (2007) 211 – 259.

[12] N. Balakrishnan, A. Childs, B. Chandrasekar, An efficient computational method for moments of order statistics under progressive censoring, Stat. Prob. Letters, 60 (2002) 359 – 365.

[13] M. Z. Raqab, N. Balakrishnan, Prediction intervals for future records, Stat. Prob. Letters, 78 (2008) 1955 – 1963.

[14] D. Krewski, Distribution-free confidence intervals for quantile intervals, J. American Stat. Association, 71 (1976) 420 – 422.

[15] E.K. Al-Hussaini, Generalized order statistics: prospective and applications, J. Appl. Stat. Sci. 4 (2004) 1 – 12.

[16] E. Cramer, Contributions to generalized order statistics, Habilitationsschrift reprint, University of Oldenburg (2003).

- [17] U. Kamps, A concept of generalized order statistics, Teubner, Stuttgart (1995).
- [18] U. Kamps, E. Cramer, On distributions of generalized order statistics, *Statistics* 35 (2001) 269 – 280.
- [19] J.F. Lawless, *Statistical model & methods for lifetime data*, Wiley, New York (1982).
- [20] W.B. Nelson, *Applied life data analysis*, Wiley, New York (1982).