

Predicting depth value of the future depth-based multivariate record

Samaneh Tat^a, Mohammad Reza Faridrohani^{1,a}

^aDepartment of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran

Abstract

The prediction problem of univariate records, though not addressed in multivariate records, has been discussed by many authors based on records values. There are various definitions for multivariate records among which depth-based records have been selected for the aim of this paper. In this paper, by means of the maximum likelihood and conditional median methods, point and interval predictions of depth values which are related to the future depth-based multivariate records are considered on the basis of the observed ones. The observations derived from some elements of the elliptical distributions are the main reason of studying this problem. Finally, the satisfactory performance of the prediction methods is illustrated via some simulation studies and a real dataset about Kermanshah city drought.

Keywords: multivariate record, depth function, depth-based multivariate record, prediction, elliptical distribution

1. Introduction

Let X_1, X_2, \dots is a sequence of independent identically distributed (i.i.d) random variables with cumulative distribution function (cdf) $F(x, \theta)$ and probability density function (pdf) $f(x, \theta)$. An observation X_j is defined as an upper record value if $X_j > X_i$ for every $j > i$. Lower records can be defined in an analogous way, only by reversing the inequalities. Records of i.i.d random variables have a unique and unambiguous definition. Univariate records, their characteristics and their applications are well described in the literature. See Resnick (1973), Arnold *et al.* (2011), Nevzorov (2001) and the references therein. A survey of literature demonstrates comprehensive studies about estimation and prediction based on the univariate records. Ahsanullah (1980), Smith (1988), Berred (1998), Raqab *et al.* (2007), Ahmadi *et al.* (2009) and Ahsanullah and Nevzorov (2015) are some examples of studies about inferential problems according to the univariate records.

Introducing the notion of record in multi-dimensional space is not as simple as uni-dimensional space. This is due to the lack of some obvious ordering properties of univariate samples in the case of multivariate observations. Hence, various definitions of multivariate records have been presented. For example, Pareto record, dominating record and chain record presented by Gnedin (2007) and Hwang and Tsai (2010) and depth-based records presented by Tat and Faridrohani (2021) are some types of multivariate records.

¹Corresponding author: Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran 1983969411, Iran. E-mail: m.faridrohani@sbu.ac.ir

Despite different definitions for multivariate records, the problem of prediction of future records has not been studied so far. The goal of this paper is to get into this problem through employing the depth-based records. In the depth-based procedure, the position of each observation is allocated by its depth value with respect to a global data cloud $\chi_s = \{\mathbf{X}_1, \dots, \mathbf{X}_s\}$. So, this value determines whether a multivariate observation can be a record. In other words, an observation is a depth-based record if its depth value is the smallest one among all the previous observations. Therefore, measuring the depth of observations in a global data cloud is an essential step in depth-based record recognition. Hence, predicting the depth value of future records can be essentially important.

The depth-based multivariate records have been defined and investigated thoroughly by Tat and Faridrohani (2021). Also, the marginal and joint distribution of depth-based record times and record values have been presented. Such information led to the problem of maximum likelihood estimation of the parameters in elliptical distribution family. This information which is useful in predicting the records is exactly what we are looking for in this paper.

Let $\mathbf{X}_1, \mathbf{X}_2, \dots$ be a sequence of multivariate observations with cdf $F(\mathbf{x}, \theta)$ and pdf $f(\mathbf{x}, \theta)$. Let $\mathbf{R}^m = \{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_m\}$ is the set of the first m depth-based records and $\mathbf{Z}^m = \{Z_1, Z_2, \dots, Z_m\}$ is the set of their corresponding depth values. Predicting \mathbf{R}_t with $t > m$ condition to the observed \mathbf{R}^m and \mathbf{Z}^m , can be an interesting problem in the field of multivariate records. Therefore, the ultimate goal is to predict the future depth-based record value, \mathbf{R}_t , $t > m$. Each record value, \mathbf{R}_i is considered as the concomitant of its depth value, Z_i . So, predicting depth value related to the future depth-based records is an underlying and primitive step in the prediction of depth-based record values. In this paper, we are concerned with the prediction of the value of Z_t , $t > m$, provided the knowledge of \mathbf{R}^m and \mathbf{Z}^m . Undoubtedly, this problem brings us one step closer to the problem of predicting the future multivariate record value, \mathbf{R}_t , $t > m$.

This paper is organized as follows. Section 2 firstly introduces depth notions and depth-based records and secondly presents the concomitant approach through which the definition of depth-based record is reconsidered and finally studies the distribution of depth value under 4 states while the observations are from one of the multivariate normal or multivariate t distributions and the records are recognized through Mahalanobis or projection depth functions. In Section 3, we introduce two methods, maximum likelihood and conditional median procedures for point and interval prediction of depth values. The aim of this paper is followed by these methods. Maximum likelihood equations don't have closed forms and a numerical procedure is needed to find maximum likelihood predictor (MLP), while conditional median predictor (CMP) can be presented in a closed form. The performance of two predictors is evaluated in Section 4 by means of the bias and mean square prediction error. The corresponding prediction interval along with simulations and an illustrative example is also discussed in this section. Finally, Section 5 includes the concluding remarks.

2. Preliminaries

As pointed out in Tat and Faridrohani (2021), in depth-based viewpoint, an observation is a proper candidate for record if it is relatively far from the previous observations. So we should be able to determine the position of each observation relative to the dataset from which the observation is derived. The notion of depth can be a good tool for this matter. In this section, at the outset, we give a brief review of the notion of data depth and present two depth functions which will be used in this paper. Then, we introduce the depth-based multivariate records and review some of their needful characteristics.

2.1. Depth notion

Data depth is a device for measuring the centrality of a given point with respect to a multivariate dataset or distribution. Employing the depth values, a center-outward ordering of points and subsequently, center-outward ranking are formed. The center-outward ranking has been widely applied in multivariate nonparametric inferences. It would be beneficial in presenting depth-based multivariate records, too.

Let \mathcal{F} be the class of distributions on \mathbb{R}^p and let $F \in \mathcal{F}$. An associated depth function $D(\mathbf{X}, F)$ is defined to provide a center-outward ordering of points $\mathbf{X} \in \mathbb{R}^p$ relative to distribution F . Based on center-outward ordering interpretation, the set of points that globally maximize depth is the center and the points near the center have higher depth Zuo and Serfling (2000a). A formal definition of statistical depth function has been presented by Zuo and Serfling (2000a) as follow:

Let, $D(\cdot, \cdot) : \mathbb{R}^p * \mathcal{F} \rightarrow \mathbb{R}^1$ is bounded, nonnegative and satisfies the following properties :

- Affine invariance: $D(A\mathbf{X} + b, F_{A\mathbf{X}+b}) = D(\mathbf{X}, F_{\mathbf{X}})$, holds for any random \mathbf{X} in \mathbb{R}^p and $p \times p$ nonsingular matrix A and p -vector b .
- Maximality at center: $D(\theta, F) = \sup_{\mathbf{x} \in \mathbb{R}^p} D(\mathbf{x}, F)$ holds if F is symmetric about θ in some sense.
- Monotonicity relative to the deepest point: for any F having deepest point θ , $D(\mathbf{x}, F) \leq D(\theta + \alpha(\mathbf{x} - \theta), F)$ holds for $\alpha \in [0, 1]$.
- Vanishing at infinity: if $\|\mathbf{x}\| \rightarrow \infty$ then $D(\mathbf{x}, F) \rightarrow 0$.

Then $D(\cdot, F)$ is called statistical depth function. For more details see Zuo and Serfling (2000a) and Serfling (2006).

$D(\mathbf{X}_j, F)$ is a measure of depth value of the observation \mathbf{X}_j with respect to the underlying distribution. When F is unknown, $D(\mathbf{X}_j, F)$ can be replaced with its sample version, $D(\mathbf{X}_j, F_s)$, which calculates the depth value of the point \mathbf{X}_j with respect to a corresponding global data cloud, $\chi_s = \{\mathbf{X}_1, \dots, \mathbf{X}_s\}$.

Hereafter, $D(\mathbf{X}_j) = D(\mathbf{X}_j, F)$ and $D_s(\mathbf{X}_j) = D(\mathbf{X}_j, F_s) = D(\mathbf{X}_j, \chi_s)$.

The word of depth was used for the first time by Tukey (1975) to introduce the halfspace depth function. After that, various depth functions were introduced and applied for measuring the depth value of observations. Here, only Mahalanobis and Projection-based depth functions are presented.

Mahalanobis depth (Liu and Singh, 1993) : The Mahalanobis depth of $\mathbf{x} \in \mathbb{R}^p$ with respect to the underlying distribution F is measured by:

$$\text{MD}(\mathbf{x}, F) = \frac{1}{1 + (\mathbf{x} - \mu_F)' \Sigma_F^{-1} (\mathbf{x} - \mu_F)},$$

where $(\mathbf{x} - \mu_F)' \Sigma_F^{-1} (\mathbf{x} - \mu_F)$ is a mahalanobis distance between \mathbf{x} and the centered vector, μ_F , with respect to the dispersion matrix of distribution F , Σ_F .

Projection-based depth (Liu, 1992; Zuo, 2003) : Suppose,

$$O(\mathbf{x}, F) = \sup_{\|u\|=1} \frac{|u' \mathbf{x} - \mu(F_u)|}{\sigma(F_u)},$$

and F_u is the distribution of $u' \mathbf{X}$, $(|u' \mathbf{x} - \mu(F_u)|)/\sigma(F_u)$ equals with 0, if $u' \mathbf{x} - \mu(F_u) = \sigma(F_u) = 0$. Then the projection-based depth (PD) of a point $\mathbf{x} \in \mathbb{R}^p$ with respect to the given F is defined as:

$$\text{PD}(\mathbf{x}, F) = \frac{1}{1 + O(\mathbf{x}, F)}.$$

Remark 1. In situations where F is unknown, the sample versions of the above mentioned depth functions are substituted. The sample version of Mahalanobis depth function denoted by $MD(\mathbf{x}, F_s) = MD(\mathbf{x}, \chi_s)$, where s is the number of observations in a global data cloud, is created when (μ, Σ) is substituted by $(\bar{\mathbf{x}}, S)$, where $\bar{\mathbf{x}}$ is the centroid of the data and S is the empirical covariance matrix with respect to χ_s . Also, the sample version of PD is denoted by $PD(\mathbf{x}, F_s) = PD(\mathbf{x}, \chi_s)$. A specific pair of (μ, σ) results in a specific PD, but the pair $(\mu, \sigma) = (\text{Med}, \text{MAD})$ have had long-term use, Med is median and MAD is median absolute deviation of χ_s .

2.2. Depth-based records

Let $\chi_s = \{\mathbf{X}_1, \dots, \mathbf{X}_s\}$ be the global data cloud, containing all the first s 'th random samples from the p -dimensional distribution function $F(\mathbf{x}, \theta)$. The observation \mathbf{X}_j is a sample depth-based record if either $j = 1$ or $j > 1$ and

$$D_s(\mathbf{X}_j) = \min(D_s(\mathbf{X}_1), D_s(\mathbf{X}_2), \dots, D_s(\mathbf{X}_j)),$$

where $D_s(\mathbf{X}_j)$ is the depth value of the observation \mathbf{X}_j with respect to the data cloud χ_s .

Also, the sequences of depth-based record times $\{T_n\}$ and record values $\{\mathbf{R}_n\}$ are defined as follows:

$$T_1 = 1, \quad \mathbf{R}_1 = \mathbf{X}_1,$$

$$T_m = \min\{j > T_{m-1} : D_s(\mathbf{X}_j) = \min(D_s(\mathbf{X}_1), D_s(\mathbf{X}_2), \dots, D_s(\mathbf{X}_j))\},$$

$$\mathbf{R}_m = \mathbf{X}_{T_m}.$$

For more familiarity with depth-based record and details on its characteristics see Tat and Faridrohani (2021).

2.3. Distributions

To get the result, it is important to know the pdf of depths related to depth-based records.

Let $\chi_s = \{\mathbf{X}_1, \dots, \mathbf{X}_s\}$ is a random sample from absolutely continuous distribution function $F_{\mathbf{X}}(\mathbf{x}, \theta)$. Correspondingly, there is a set of depth values $\mathcal{D}_s = \{D_s(\mathbf{X}_1), \dots, D_s(\mathbf{X}_s)\}$, $s \geq 1$, where $Y_j = D_s(\mathbf{X}_j)$ is a univariate random variable with continuous distribution function $F_Y(y)$. \mathcal{D}_s is sorted in a descending order. Sorting is done by applying the center-outward ordering scheme such that the last member of the sorted \mathcal{D}_s is a lower record in the sense of univariate records. In this way, pairs of order statistics $(\mathbf{X}_{[1]}, D_s(\mathbf{X}_{[1]})), \dots, (\mathbf{X}_{[s]}, D_s(\mathbf{X}_{[s]}))$ are achieved. Employing the concomitant procedure and according to the definition of depth-based records, \mathbf{X}_j is a depth-based record if $D_s(\mathbf{X}_j)$ is a lower record in the set \mathcal{D}_j .

Now suppose that $\mathbf{R}_j = \mathbf{X}_{T_j}$, $j = 1, 2, \dots, m$, then $\mathbf{r} = \{\mathbf{r}_1, \dots, \mathbf{r}_m\}$ denotes a set of the first m observed depth-based record values from χ_s . Correspondingly, $\mathbf{z} = \{z_1, z_2, \dots, z_m\}$ are the depth values of records, $z_j = D_s(\mathbf{r}_j)$ for $j = 1, \dots, m$, and $z_1 > z_2 > \dots > z_m$.

According to Tat and Faridrohani (2021), the joint pdf of the first m depth of depth-based records, $\mathbf{Z}^m = \{Z_1, Z_2, \dots, Z_m\}$, is similar to its ordinary univariate records obtained by Ahsanullah (2009) as follows:

$$f_{Z_1, \dots, Z_m}(z_1, \dots, z_m) = \frac{f_Y(z_1)}{F_Y(z_1)} \frac{f_Y(z_2)}{F_Y(z_2)} \dots \frac{f_Y(z_{m-1})}{F_Y(z_{m-1})} f_Y(z_m), \quad (2.1)$$

The marginal pdf of depth of the j^{th} depth-based record, Z_j , is given by

$$f_{Z_j}(z) = \frac{f_Y(z) (-\log F_Y(z))^{j-1}}{(j-1)!}, \quad (2.2)$$

Also the conditional pdf of Z_t given Z_m , $m < t < s$, is

$$f_{t|m}(z_t | z_m) = \frac{f_Y(z_t) (\log F_Y(z_m) - \log F_Y(z_t))^{t-m-1}}{F_Y(z_m)(t-m-1)!}. \quad (2.3)$$

It should be noted that $F_Y(y)$ depends on $F_{\mathbf{X}}(\mathbf{x}, \theta)$ and cannot be independently determined. In the next subsection, we will explain their relevancy.

Now we intend to explicate the nexus between the distribution of the multivariate observations and the distribution of their depth values.

We understand from the literature that if observations are from elliptical distribution and the depth functions have some property, then the distribution of the depth values has a known form.

A random vector \mathbf{X} in \mathbb{R}^p has an elliptical distribution, $\mathbf{X} \sim E_d(h; \boldsymbol{\mu}, \boldsymbol{\Sigma})$, if its density function is of the form

$$f(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = C_p |\boldsymbol{\Sigma}|^{-\frac{1}{2}} h\left((\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right), \quad (2.4)$$

where C_p is a constant depending on p and the function h ; $\boldsymbol{\mu}$ is the center of the distribution; $\boldsymbol{\Sigma}$ is a positive definite matrix and $h(\cdot)$ is a nonnegative function (Liu *et al.*, 1999).

According to Zuo and Serfling (2000b), suppose that \mathbf{X} has elliptical distribution defined in (2.4) and $D(\cdot, F)$ is affine invariant (property P1) and attains maximum value at $\boldsymbol{\mu}$ (property P2), then $D(\mathbf{x}, F)$ can be written in the form of

$$D(\mathbf{x}, F) = g\left((\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right), \quad (2.5)$$

for some nonincreasing function g .

The function g is accessible for both Mahalanobis and projection depth functions. For the Mahalanobis depth function and $t \in \mathbb{R}$, $g(t) = 1/(1+t)$ and based on Zuo (2003), $g(t) = 1/(1 + \sqrt{t})$, for the projection depth function.

The family of multivariate t distribution and multivariate normal distribution are members of the elliptical distribution and can be written in the form of (2.5) easily.

If we suppose the multivariate observations are from multivariate t or multivariate normal distributions and the applied depth function is Mahalanobis or projection, we will have four states which will be presented in continue.

S1. Assume $\mathbf{X}_1, \mathbf{X}_2, \dots$ be a sequence of i.i.d observations from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Suppose that Mahalanobis depth function is applied for measuring the depth values of the observations. We define, $Y = D(\mathbf{x}) = 1/(1+T)$, where $T = (\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_p^2$, a Chi-square distribution with p degrees of freedom (DeGroot, 2005). Then the distribution function of Y is

$$F_Y(y) = \int_{\frac{1-y}{y}}^{\infty} \frac{1}{\Gamma(p/2)} \left(\frac{1}{2}\right)^{\frac{p}{2}} t^{\frac{p}{2}-1} e^{-\frac{t}{2}} dt. \quad (2.6)$$

- S2. Assume $\mathbf{X}_1, \mathbf{X}_2, \dots$ be a sequence of i.i.d observations from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Suppose that projection depth function is employed for measuring the depth values of the observations. Then $Y = D(\mathbf{X}) = 1/(1 + \sqrt{T})$ where $T = (\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu})$. The distribution function of Y is of the form

$$F_Y(y) = \int_{\left(\frac{1-y}{y}\right)^2}^{\infty} \frac{1}{\Gamma(p/2)} \left(\frac{1}{2}\right)^{\frac{p}{2}} t^{\frac{p}{2}-1} e^{-\frac{t}{2}} dt. \quad (2.7)$$

- S3. Suppose $\mathbf{X}_1, \mathbf{X}_2, \dots$ be a sequence of i.i.d observations from $t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$. If Mahalanobis depth function is used for calculating the depth values of the observations, then $Y = D(\mathbf{x}) = 1/(1 + pF)$ where $F = T/p = (1/p)(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim F(p, \nu)$, the F distribution with p numerator degrees of freedom and ν denominator degrees of freedom. The distribution function of Y is given by

$$F_Y(y) = \int_{\frac{1-y}{py}}^{\infty} \frac{\Gamma((p+\nu)/2)}{\Gamma(\nu/2)\Gamma(p/2)} \left(\frac{p}{\nu}\right)^{\frac{p}{2}} \frac{f^{\frac{p-2}{2}}}{[1 + (fp/\nu)]^{\frac{p+\nu}{2}}} df. \quad (2.8)$$

- S4. Let $\mathbf{X}_1, \mathbf{X}_2, \dots$ be a sequence of i.i.d observations from $t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$. Suppose depth values of observations are calculated using projection function. So, $Y = D(\mathbf{x}) = 1/(1 + \sqrt{pF})$, where $F = T/p = (1/p)(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim F(p, \nu)$, has the distribution function of the form

$$F_Y(y) = \int_{\frac{(1-y)^2}{py^2}}^{\infty} \frac{\Gamma((p+\nu)/2)}{\Gamma(\nu/2)\Gamma(p/2)} \left(\frac{p}{\nu}\right)^{\frac{p}{2}} \frac{f^{\frac{p-2}{2}}}{[1 + (fp/\nu)]^{\frac{p+\nu}{2}}} df. \quad (2.9)$$

Prediction depth of the future depth-based records will be done under each of the above states in what follows.

3. Prediction

In this section, we deal with the problem of predicting depth of the depth-based records. Suppose that we observe the first m depth-based records, $\mathbf{r} = \{\mathbf{r}_1, \dots, \mathbf{r}_m\}$, and the corresponding depth values, $\mathbf{z} = \{z_1, z_2, \dots, z_m\}$, respectively from $F_{\mathbf{X}}(\mathbf{x}, \theta)$ and $F_Y(y)$, where θ is unknown. Our aim is to predict Z_t , $t > m$, having observed \mathbf{z} with two different schemes; maximum likelihood and conditional median prediction.

3.1. Maximum likelihood prediction

In this subsection, our objective is to predict the depth of the future depth-based records using maximum likelihood approach. The joint predictive likelihood function of Z_t is given as

$$\begin{aligned} L(\mathbf{z}, Z_t) &= f(\mathbf{z})h(z_t | \mathbf{z}) \\ &= f(\mathbf{z})f_{t|m}(z_t | z_m), \end{aligned}$$

where the last equality is obtained by the Markovian property of univariate records. If there exists $Z_t^0 = Z_t^{mle}$ such that

$$L(\mathbf{z}, Z_t^0) = \sup_{z_t} L(\mathbf{z}, Z_t),$$

then Z_t^0 is the maximum likelihood predictor (MLP) of Z_t .

Using the equations (2.1)–(2.3), the predictive likelihood function is

$$L(\mathbf{z}, Z_t) = \prod_{j=1}^m \frac{f_Y(z_j)}{F_Y(z_j)} \times \frac{f_Y(z_t) (\log F_Y(z_m) - \log F_Y(z_t))^{t-m-1}}{(t-m-1)!}, \quad (3.1)$$

and consequently, the predictive log-likelihood function is given by

$$l(\mathbf{z}, Z_t) = \sum_{j=1}^m (\log f_Y(z_j) - \log F_Y(z_j)) + (t-m-1) \log (\log F_Y(z_m) - \log F_Y(z_t)) + \log f_Y(z_t) - \log (t-m-1)!,$$

The MLP of Z_t can be obtained from the following normal equation.

$$0 = \frac{\partial l(\mathbf{z}, Z_t)}{\partial Z_t} = \frac{f'_Y(z_t)}{F_Y(z_t)} - (t-m-1) \frac{f_Y(z_t)}{F_Y(z_t) \times (\log F_Y(z_m) - \log F_Y(z_t))}, \quad (3.2)$$

The above equations cannot be solved in closed form. Thus, we must use a numerical procedure to find the maximum likelihood predictor (MLP).

3.2. Conditional median prediction

Using the conditional distribution of Z_t given Z_m is an approach for prediction the future records. Here, we use this method for predicting depth of the future depth-based records.

Under the assumptions mentioned in this section and using the Markovian property of univariate records, the conditional density of Z_t given \mathbf{z} is as

$$f(z_t | \mathbf{z}) = f(z_t | z_m) = \frac{f_Y(z_t)}{F_Y(z_m)} \times \frac{(\log F_Y(z_m) - \log F_Y(z_t))^{t-m-1}}{(t-m-1)!}, \quad z_t < z_m, \quad (3.3)$$

The median of this distribution is called the conditional median predictor (CMP). The CMP is a function of z_m . So assume that

$$\hat{Z}_t^{CMP} = h(z_m) = \omega.$$

Hence,

$$P(h(z_m) \leq Y | z_m) = \frac{1}{2}.$$

Using the equation (3.3)

$$\int_{\omega}^{z_m} \frac{f_Y(z_t)}{F_Y(z_m)} \times \frac{(\log F_Y(z_m) - \log F_Y(z_t))^{s-m-1}}{(s-m-1)!} dz_t = \frac{1}{2}.$$

If we assume that $Z_m = z_m$ has been observed, then by taking $q = \log F_Y(z_m) - \log F_Y(z_t)$, the above equality can be rewritten as

$$\int_0^{\log \frac{F_Y(z_m)}{F_Y(\omega)}} \frac{q^{t-m-1}}{(t-m-1)!} e^{-q} dq = \frac{1}{2}.$$

Table 1: Bias and MSPE of MLP and CMP for multivariate normal distribution with parameter sets (4.1) and (4.2)

Distributions	m	t	Mahalanobis				Projection			
			MLP		CMP		MLP		CMP	
			Bias	MSPE	Bias	MSPE	Bias	MSPE	Bias	MSPE
Multivariate normal distribution with parameter set (4.1)	6	7	0.0501	0.0055	0.0030	2.2×10^{-4}	0.0789	0.0111	-0.0144	5.8×10^{-4}
	6	8	0.0636	0.0072	0.0039	2.2×10^{-4}	0.0731	0.0103	-0.0260	1.2×10^{-3}
	6	9	0.0738	0.0087	0.0040	2.4×10^{-4}	0.0871	0.0122	-0.0342	1.7×10^{-3}
	6	10	0.0833	0.0104	0.0044	2.5×10^{-4}	0.0903	0.0132	-0.0420	2.3×10^{-3}
	7	8	0.0421	0.0042	0.0009	1.2×10^{-4}	0.0725	0.0093	-0.0127	3.9×10^{-4}
	7	9	0.0523	0.0052	0.0010	1.7×10^{-4}	0.0698	0.0080	-0.0216	7.5×10^{-4}
	7	10	0.0618	0.0064	-0.0015	1.8×10^{-4}	0.0816	0.0105	-0.0300	1.2×10^{-3}
	8	9	0.0345	0.0031	-0.0011	9.7×10^{-5}	0.0677	0.0074	-0.0096	2.1×10^{-4}
	8	10	0.0440	0.0040	-0.0007	1.0×10^{-4}	0.0700	0.0076	-0.0184	5.2×10^{-4}
	9	10	0.0291	0.0025	-0.0024	6.9×10^{-4}	0.0642	0.0067	-0.0091	1.7×10^{-4}
Multivariate normal distribution with parameter set (4.2)	6	7	0.0545	0.0058	-0.0005	2.9×10^{-4}	0.0679	0.0094	0.0054	4.4×10^{-4}
	6	8	0.0705	0.0078	0.0031	3.3×10^{-4}	0.0691	0.0099	0.0080	5.9×10^{-4}
	6	9	0.0794	0.0090	0.0039	3.4×10^{-4}	0.0784	0.0114	0.0076	6.3×10^{-4}
	6	10	0.0932	0.0119	0.0041	4.2×10^{-4}	0.0792	0.0116	0.0040	5.2×10^{-4}
	7	8	0.0494	0.0047	-0.0007	1.6×10^{-4}	0.0585	0.0069	0.0015	2.8×10^{-4}
	7	9	0.0610	0.0062	0.0012	1.6×10^{-4}	0.0603	0.0073	0.0015	2.9×10^{-4}
	7	10	0.0717	0.0079	0.0017	1.7×10^{-4}	0.0698	0.0088	0.0017	3.0×10^{-4}
	8	9	0.0401	0.0037	-0.0004	9.7×10^{-5}	0.0570	0.0061	0.0013	2.1×10^{-4}
	8	10	0.0503	0.0044	-0.0005	1.1×10^{-4}	0.0654	0.0089	0.0043	2.2×10^{-4}
	9	10	0.0337	0.0025	-0.0017	6.7×10^{-5}	0.0599	0.0061	0.0044	1.3×10^{-4}

So,

$$\hat{Z}_t^{CMP} = F_Y^{-1} \left(F_Y(z_m) \times e^{-Med(W)} \right), \tag{3.4}$$

where $W \sim G(t-m, 1)$ has gamma distribution with the shape parameter $(t-m)$ and the scale parameter 1. Also, $Med(W)$ stands for the median of W .

Suppose that $M = \log(F_Y(z_m))/(F_Y(z_t))$. It is easily shown $2M|z_m \sim \chi_{2(t-m)}^2$. Thus, the exact $(1 - \gamma)100\%$ prediction interval for Z_t is

$$I = \left(F^{-1} \left(\exp \left\{ \frac{-1}{2} \chi_{2(t-m), (1-\gamma/2)}^2 \right\} F_Y(z_m) \right), F^{-1} \left(\exp \left\{ \frac{-1}{2} \chi_{2(t-m), (\gamma/2)}^2 \right\} F_Y(z_m) \right) \right), \tag{3.5}$$

where $\chi_{r,p}^2$ is $100p^{th}$ percentile from the χ^2 distribution with degrees of freedom r .

4. Simulation and data analysis

In this section, we discuss some predictions of depth values related to the future depth-based record data extracted from a practical dataset by means of multivariate normal distribution and then conduct some simulation studies to assess the performance of MLP of depth value due to the future depth-based record as well as its CMP. For this purpose, we employ multivariate normal and multivariate t distributions along with the Mahalanobis and projection-based depth functions. All the computations are performed using R Software.

4.1. Simulation results

Here, we evaluate the performance of the different methods of predicting depth of the t^{th} future depth-based records condition to observe the first m record values. Precisely, we evaluate the performance of the MLP and CMP in terms of bias and MSPE.

Table 2: Bias and MSPE of MLP and CMP for multivariate t distribution with 3 degrees of freedom and parameter sets (4.1) and (4.2)

Distributions	m	t	Mahalanobis				Projection			
			MLP		CMP		MLP		CMP	
			Bias	MSPE	Bias	MSPE	Bias	MSPE	Bias	MSPE
Multivariate t distribution with parameter set (4.1)	6	7	0.0331	0.0043	0.0042	3.1×10^{-4}	-0.0015	0.0047	0.0055	4.6×10^{-4}
	6	8	0.0484	0.0060	0.0072	3.1×10^{-4}	0.0153	0.0048	0.0115	5.8×10^{-4}
	6	9	0.0580	0.0073	0.0074	3.9×10^{-4}	0.0280	0.0051	0.0144	5.8×10^{-4}
	6	10	0.0638	0.0083	0.0051	4.2×10^{-4}	0.0412	0.0063	0.0152	6.1×10^{-4}
	7	8	0.0222	0.0033	0.0008	2.0×10^{-4}	0.0168	0.0041	0.0019	2.3×10^{-4}
	7	9	0.0319	0.0042	0.0032	1.4×10^{-4}	0.0315	0.0045	0.0071	3.3×10^{-4}
	7	10	0.0375	0.0049	0.0035	1.5×10^{-4}	0.0459	0.0051	0.0073	3.3×10^{-4}
	8	9	0.0103	0.0026	0.0001	6.7×10^{-5}	0.0305	0.0036	0.0011	1.8×10^{-4}
	8	10	0.0160	0.0030	0.0020	6.9×10^{-5}	0.0442	0.0045	0.0022	2.1×10^{-4}
	9	10	-0.0018	0.0022	0.0002	1.5×10^{-5}	0.0295	0.0033	-0.0024	1.3×10^{-4}
Multivariate t distribution with parameter set (4.2)	6	7	0.0375	0.0046	0.0036	3.1×10^{-4}	0.0395	0.0049	0.0051	4.5×10^{-4}
	6	8	0.0531	0.0069	0.0080	3.0×10^{-4}	0.0535	0.0070	0.0135	6.9×10^{-4}
	6	9	0.0609	0.0076	0.0082	3.2×10^{-4}	0.0635	0.0086	0.0162	7.1×10^{-4}
	6	10	0.0686	0.0091	0.0084	3.3×10^{-4}	0.0637	0.0086	0.0160	7.2×10^{-4}
	7	8	0.0243	0.0035	0.0021	1.5×10^{-4}	0.0248	0.0034	0.0042	3.8×10^{-4}
	7	9	0.0341	0.0043	0.0043	1.5×10^{-4}	0.0342	0.0043	0.0092	4.6×10^{-4}
	7	10	0.0411	0.0050	0.0049	1.6×10^{-4}	0.0405	0.0049	0.0098	5.0×10^{-4}
	8	9	0.0127	0.0027	0.0004	6.0×10^{-5}	0.0125	0.0026	0.0013	1.7×10^{-4}
	8	10	0.0194	0.0030	0.0020	6.7×10^{-5}	0.0201	0.0033	0.0038	2.4×10^{-4}
	9	10	7.3×10^{-5}	0.0019	-7.1×10^{-6}	2.9×10^{-5}	0.0021	0.0011	-0.0011	1.4×10^{-4}

Table 3: Bias and MSPE of MLP and CMP for multivariate t distribution with 10 degrees of freedom and parameter sets (4.1) and (4.2)

Distributions	m	t	Mahalanobis				Projection			
			MLP		CMP		MLP		CMP	
			Bias	MSPE	Bias	MSPE	Bias	MSPE	Bias	MSPE
Multivariate t distribution with parameter set (4.1)	6	7	0.0138	0.0031	9.6×10^{-4}	3.2×10^{-4}	0.0199	0.0040	0.0078	4.6×10^{-4}
	6	8	0.0291	0.0041	0.0060	3.9×10^{-4}	0.0337	0.0049	0.0140	8.6×10^{-4}
	6	9	0.0429	0.0052	0.0068	4.2×10^{-4}	0.0492	0.0065	0.0170	8.7×10^{-4}
	6	10	0.0511	0.0060	0.0070	5.1×10^{-4}	0.0550	0.0071	0.0186	9.0×10^{-4}
	7	8	0.0058	0.0025	3.8×10^{-4}	2.2×10^{-4}	0.0099	0.0029	0.0033	4.3×10^{-4}
	7	9	0.0195	0.0030	0.0019	2.4×10^{-4}	0.0219	0.0035	0.0074	4.5×10^{-4}
	7	10	0.0277	0.0034	0.0043	2.4×10^{-4}	0.0315	0.0041	0.0100	5.0×10^{-4}
	8	9	0.0042	0.0021	-0.0022	1.5×10^{-4}	0.0022	0.0023	8.9×10^{-4}	2.3×10^{-4}
	8	10	0.0083	0.0023	0.0014	1.7×10^{-4}	0.0104	0.0024	0.0041	2.6×10^{-4}
	9	10	-0.0013	0.0017	1.2×10^{-4}	6.3×10^{-5}	-0.0070	0.0018	-1.9×10^{-5}	1.4×10^{-4}
Multivariate t distribution with parameter set (4.2)	6	7	0.0104	0.0032	0.0012	3.2×10^{-4}	0.0024	0.0045	0.0083	8.6×10^{-4}
	6	8	0.0249	0.0042	0.0076	4.1×10^{-4}	0.0066	0.0046	0.0144	9.3×10^{-4}
	6	9	0.0350	0.0047	0.0084	4.5×10^{-4}	0.0173	0.0047	0.0165	9.4×10^{-4}
	6	10	0.0460	0.0054	0.0085	4.8×10^{-4}	0.0217	0.0051	0.0167	9.6×10^{-4}
	7	8	0.0011	0.0026	0.0017	1.8×10^{-4}	0.0036	0.0036	0.0032	4.5×10^{-4}
	7	9	0.0139	0.0028	0.0042	2.2×10^{-4}	0.0180	0.0037	0.0065	4.8×10^{-4}
	7	10	0.0237	0.0031	0.0050	2.3×10^{-4}	0.0277	0.0038	0.0068	4.9×10^{-4}
	8	9	-0.0060	0.0021	9.3×10^{-4}	1.2×10^{-4}	0.0125	0.0032	-6.1×10^{-5}	2.1×10^{-4}
	8	10	0.0043	0.0023	0.0011	1.3×10^{-4}	0.0267	0.0033	-1.1×10^{-4}	3.0×10^{-4}
	9	10	-0.0010	0.0018	-1.3×10^{-4}	7.6×10^{-5}	0.0110	0.0027	-3.4×10^{-4}	1.6×10^{-4}

For data simulation, multivariate normal, $N_p(\mu, \Sigma)$ and multivariate t distributions, $t_p(\mu, \Sigma, \nu)$, for $p = 2$ and the following two sets of parameters are employed.

$$\mu = \begin{pmatrix} 5 \\ 7 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 20 & 17 \\ 17 & 15 \end{pmatrix}, \quad \nu = 3, 10, \quad (4.1)$$

$$\mu = \begin{pmatrix} 5 \\ 7 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 20 & 10 \\ 10 & 15 \end{pmatrix}, \quad \nu = 3, 10. \quad (4.2)$$

Table 4: ALs of PIs for multivariate normal and t distributions under parameter sets (4.1) and (4.2)

Depth function	m	t	Parameter set (4.1)			Parameter set (4.2)		
			Normal	t with $\nu = 3$	t with $\nu = 10$	Normal	t with $\nu = 3$	t with $\nu = 10$
Mahalanobis depth function	6	7	(0.0603,0.1122)	(0.0045,0.0598)	(0.0306,0.0952)	(0.0608,0.1138)	(0.0044,0.0584)	(0.0307,0.0954)
	6	8	(0.0489,0.1065)	(0.0013,0.0505)	(0.0192,0.0874)	(0.0493,0.1080)	(0.0012,0.0495)	(0.0193,0.0877)
	6	9	(0.0420,0.0981)	(0.0004,0.0380)	(0.0131,0.0761)	(0.0423,0.0994)	(0.0004,0.0373)	(0.0132,0.0763)
	6	10	(0.0372,0.0893)	(0.0001,0.0269)	(0.0093,0.0647)	(0.0374,0.0904)	(0.0001,0.0265)	(0.0094,0.0650)
	7	8	(0.0557,0.0965)	(0.0031,0.0394)	(0.0257,0.0744)	(0.0557,0.0969)	(0.0030,0.0383)	(0.0259,0.0749)
	7	9	(0.0459,0.0924)	(0.0009,0.0335)	(0.0164,0.0689)	(0.0459,0.0927)	(0.0008,0.0325)	(0.0165,0.0694)
	7	10	(0.0398,0.0860)	(0.0003,0.0254)	(0.0113,0.0608)	(0.0398,0.0863)	(0.0003,0.0247)	(0.0113,0.0612)
	8	9	(0.0509,0.0831)	(0.0020,0.0245)	(0.0215,0.0592)	(0.0512,0.0840)	(0.0020,0.0251)	(0.0222,0.0610)
	8	10	(0.0426,0.0800)	(0.0006,0.0209)	(0.0139,0.0552)	(0.0428,0.0809)	(0.0006,0.0214)	(0.0142,0.0569)
	9	10	(0.0469,0.0730)	(0.0012,0.0150)	(0.0175,0.0455)	(0.0471,0.0733)	(0.0013,0.0157)	(0.0184,0.0482)
Projection depth function	6	7	(0.1058,0.1128)	(0.0282,0.0904)	(0.0943,0.1387)	(0.1433,0.1594)	(0.0258,0.0913)	(0.0952,0.1402)
	6	8	(0.1029,0.1124)	(0.0153,0.0845)	(0.0781,0.1354)	(0.1369,0.1582)	(0.0154,0.0853)	(0.0788,0.1369)
	6	9	(0.1005,0.1115)	(0.0088,0.0751)	(0.0663,0.1299)	(0.1321,0.1563)	(0.0089,0.0758)	(0.0670,0.1313)
	6	10	(0.0985,0.1106)	(0.0053,0.647)	(0.0570,0.1235)	(0.1280,0.1541)	(0.0053,0.0654)	(0.0575,0.1248)
	7	8	(0.0927,0.0969)	(0.0242,0.0782)	(0.0929,0.1363)	(0.1452,0.1613)	(0.0243,0.0785)	(0.0941,0.1381)
	7	9	(0.0909,0.0967)	(0.0131,0.0730)	(0.0770,0.1330)	(0.1388,0.1602)	(0.0131,0.0733)	(0.0779,0.1349)
	7	10	(0.0893,0.0962)	(0.0076,0.0649)	(0.0654,0.1277)	(0.1338,0.1583)	(0.0076,0.0651)	(0.0662,0.1295)
	8	9	(0.0808,0.0834)	(0.0199,0.0647)	(0.0888,0.1296)	(0.1440,0.1593)	(0.0208,0.0673)	(0.0899,0.1312)
	8	10	(0.0795,0.0833)	(0.0108,0.0604)	(0.0737,0.1266)	(0.1377,0.1582)	(0.0112,0.0628)	(0.0745,0.1282)
	9	10	(0.0715,0.0733)	(0.0162,0.0530)	(0.0835,0.1213)	(0.1406,0.1547)	(0.0170,0.0553)	(0.0838,0.1218)

In the above sets, Σ is the covariance matrix for the multivariate normal distribution and the scatter matrix for the multivariate t distribution. In the two parameter sets, only the off-diagonal entries in Σ have been varied. For the parameter set (4.1), the correlation is more intense; the density is more compressed, and generally, the contours form ellipses.

In each setting, 200 repetitions of independent random sample are generated sequentially and then the m MD-based and PD-based records, $\mathbf{r} = \{\mathbf{r}_1, \dots, \mathbf{r}_m\}$ and their corresponding depth values, $\mathbf{z} = \{z_1, \dots, z_m\}$, for $m = 6, 7, 8, 9$, are extracted. Using the observed records and their depth values, point predictors of depth values are computed as well as the corresponding prediction intervals (PI) for $t > m$.

The MLP is derived from the solution of the equation (3.2) and according to equations (2.6) to (2.9). Since this equation should be solved analytically, we use the Nelder-Mead method implemented in the built-in maxLik function in the R package maxLik (Henningsen and Toomet, 2011) to maximize the likelihood functions in terms of suitable parameter. Here, the record to be predicted is known as the suitable parameter. Applying the optimization algorithm, one initial value for the parameter should be allocated. Our suggestion for the specification of this initial value is the average of the depth values related to the m observed depth-based records.

The CMP of depth values related to the t^{th} depth-based record is resulted from the equation (3.4) and its PI is derived from the interval (3.5).

The bias and MSPEs of the predictors are computed for each method over 200 replications and these are all presented in Tables 1–3. Table 4 represents the average lengths (ALs) of 95% PI for the depth value related to the future t^{th} depth-based record, based on the relation (3.5).

From the results reported in the Tables 1–3, some general facts are derived. These outcomes demonstrate proper predictions of depth values related to the future t^{th} depth-based records. It can be quickly deduced from the results that by increasing the number of observed records, the prediction value gets closer to the actual value. Even when the gap between m and t is more less, predictions are made more accurately.

It turns out that regardless of depth and distribution functions, biases and MSPEs associated with the CMP are less than MLP. Nevertheless, conditional median procedure leads to more accurate predictions. Also, depth values related to the future Mahalanobis-based records are predicted more pre-

Table 5: MD- and PD-based records in the term of TMP and AMT variables

MD-based record				PD-based record			
Record time	Year	(TMP, AMT)	Depth	Record time	Year	(TMP, AMT)	Depth
1	1951	(0.99, 2.36)	0.290	1	1951	(0.99, 2.36)	0.294
13	1963	(0.99, 2.38)	0.233	2	1952	(0.99, 2.44)	0.203
21	1971	(0.99, 2.34)	0.179	13	1963	(0.99, 2.38)	0.197
48	1998	(0.99, 2.60)	0.144	57	2007	(0.99, 2.55)	0.163
57	2007	(0.99, 2.55)	0.097	66	2016	(0.99, 2.54)	0.127
66	2016	(0.99, 2.54)	0.070	-	-	-	-

cisely than projection-based records.

Another derived conclusion from the above tables is when the Mahalanobis-based records are generated from multivariate normal distributions with parameters in (4.1) and have more intense contours, MLPs of depth values have less biases and MSPEs. Such inductions can not be deduced from other distributions and depth functions.

What can generally be concluded is that by increasing the number of the observed records, m , prediction under both procedures, maximum likelihood and conditional median, is well done.

From Table 4, it is observed that PIs based on projection-based records under multivariate normal distribution are narrower than Mahalanobis-based records. In addition, the difference between the length of Mahalanobis-based and projection-based prediction intervals is notable. This issue is quit the contrary for multivariate t distribution with $\nu = 3$. In fact, PIs related to the depth values of the future Mahalanobis-based records from t distribution with 3 degrees of freedom have shorter lengths than projection-based records.

According to the results due to PIs, there is a clear evidence that predictions of depth values from multivariate distributions with parameter set (4.1) perform same as parameter set (4.2). This is true for both Mahalanobis and projection depth functions.

4.2. Data analysis

By considering a real dataset, we illustrate the performance of the ML and CM predictors of the depth value related to the future depth-based records in practice. The dataset is about the Kermanshah city drought. It consists of two measurements, total monthly precipitation (TMP) and average monthly temperature (AMT) during 66 water years, 1951–2016. The term water year is equivalent to 12-months period for which precipitation totals are measured. In Iran, water year is defined the period between September 23th of one year to September 22th of the following year. The data used in this case are gathered by the department of meteorology of Kermanshah city, which is a subset of Islamic Republic (I.R.) of Iran meteorological organization (IRIMO) and are accessible from “www.kermanshahmet.ir”. These data have been previously analyzed earlier by Tat and Faridrohani (2021).

Since the distribution of Kermanshah city did not follow from a multivariate normal distribution, we applied a box-cox transformation to get a dataset with a multivariate normal distribution. The MD-based and PD-based records along with the corresponding depth values were extracted from this transformed dataset. The information about records is reported in the Table 5.

In view of the depth-based approach, 6 MD-based and 5 PD-based records are recognized during 66 water years for Kermanshah city. These records consist of both wetness and dryness records. For more familiarity with them see Tat and Faridrohani (2021). In order to evaluate the performance of prediction procedures in reality, we eliminated the last depth-based record and endeavored to predict

Table 6: MLP and CMP of depth value related to the t^{th} depth-based record

Depth function	m	s	Z_t	\hat{Z}_t^{MLP}	\hat{Z}_t^{CMP} (PI)
Mahalanobis	5	6	0.70	0.82	0.79 (0.565, 0.965)
Projection	4	5	0.127	0.143	0.138 (0.126, 0.163)

its depth value condition to the former records via ML and CM procedures from equations (3.2) and (3.4).

The results of predictions of depth values related to the t^{th} depth-based records condition to the first m records are reported in the Table 6. It also contains the 95% prediction interval which is calculated from the relation (3.5).

With regard to the displayed results in the above table, the CMP is closer to the actual value than the MLP for both Mahalanobis and projection depth functions. It also should be noted that the Mahalanobis depth function has made more accurate predictions than the projection depth function.

5. Concluding remarks

The problem of prediction is an important issue in the field of records. It has been studied thoroughly for univariate records, while there is no entrance to prediction issue of multivariate records. For this purpose, we selected depth-based record from different definitions of multivariate records. So this paper can be considered a step forward in the problem of multivariate record prediction. Since a depth-based record is recognized by its depth value, we decided to study the prediction problem in two parts. In the first part, we focused on the prediction of depth values related to the depth-based records which is undoubtedly an underlying step for the second part. For this purpose, we supposed the observations are from multivariate normal or multivariate t distributions and records are recognized by Mahalanobis and projection depth functions from these observations.

In this paper, we studied the problem of prediction of depth values related to the future depth-based records through maximum likelihood (ML) and conditional median (CM) procedures. Besides, we could build a prediction interval for depth values. Finally, we evaluated the performance of both procedures by some simulation studies and a real dataset about Kermanshah city drought. Both results of the simulations and real dataset demonstrated satisfactory performance for both MLP and CMP.

The second part dealing with the prediction problem of depth-based record values will be considered in the future works.

References

- Ahmadi J, Jafari Jozani M, Marchand É, and Parsian A (2009). Prediction of k-records from a general class of distributions under balanced type loss functions, *Metrika*, **70**, 19–33.
- Ahsanullah M (1980). Linear prediction of record values for the two parameter exponential distribution, *Annals of the Institute of Statistical Mathematics*, **32**, 363–368.
- Ahsanullah M (2009). Records and concomitants, *Bulletin of the Malaysian Mathematical Sciences Society*. Second Series, **32**, 101–117.
- Ahsanullah M and Nevzorov VB (2015). *Records via Probability Theory*, Atlantis Press, Paris.
- Arnold BC, Balakrishnan N, and Nagaraja HN (2011). *Records*, John Wiley & Sons, New York.
- Berred AM (1998). Prediction of record values, *Communications in Statistics-Theory and Methods*, **27**, 2221–2240.
- DeGroot MH (2005). *Optimal Statistical Decisions*, John Wiley & Sons, New York.
- Gnedin A (2007). The chain records, *Electronic Journal of Probability*, **12**, 767–786.

- Henningsen A and Toomet O (2011). maxLik: A package for maximum likelihood estimation in R, *Computational Statistics*, **26**, 443–458.
- Hwang HK and Tsai TH (2010). Multivariate records based on dominance, *Electronic Journal of Probability*, **15**, 1863–1892.
- Liu RY (1992). Data depth and multivariate rank tests. In *Proceedings of 2nd International Conference on Stat. Data Analysis Based on the L-1 Norm and Related Methods Y Dodge Ed*, North-Holland, 279–294.
- Liu RY, Parelius JM, and Singh K (1999). Multivariate analysis by data depth: Descriptive statistics, graphics and inference (with discussion and a rejoinder by liu and singh), *The Annals of Statistics*, **27**, 783–858.
- Liu RY and Singh K (1993). A quality index based on data depth and multivariate rank tests, *Journal of the American Statistical Association*, **88**, 252–260.
- Nevzorov VB (2001). *Records: Mathematical Theory*, American Mathematical Society, Providence, RI.
- Raqab MZ, Ahmadi J, and Doostparast M (2007). Statistical inference based on record data from Pareto model, *Statistics*, **41**, 105–118.
- Resnick SI (1973). Record values and maxima, *The Annals of Probability*, **1**, 650–662.
- Serfling R (2006). Depth functions in nonparametric multivariate inference, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, **72**, 1–16.
- Smith RL (1988). Forecasting records by maximum likelihood, *Journal of the American Statistical Association*, **83**, 331–338.
- Tat S and Faridrohani MR (2021). A new type of multivariate records: Depth-based records, *Statistics*, **55**, 296–320.
- Tukey J (1975). Mathematics and picturing data, *Proceedings of International Congress of Mathematicians*, **2**, 523–531.
- Zuo Y (2003). Projection-based depth functions and associated medians, *The Annals of Statistics*, **31**, 1460–1490.
- Zuo Y and Serfling R (2000a). General notions of statistical depth function, *The Annals of Statistics*, **28**, 461–482.
- Zuo Y and Serfling R (2000b). Structural properties and convergence results for contours of sample statistical depth functions, *The Annals of Statistics*, **28**, 483–499.

Received February 19, 2023; Revised February 25, 2023; Accepted February 25, 2023