Turkish Journal of Mathematics

http://journals.tubitak.gov.tr/math/

Research Article

Turk J Math
(2019) 43: 1474 – 1491
© TÜBİTAK
doi:10.3906/mat-1901-54

On distribution of upper marginal records in bivariate random sequences

Gülder KEMALBAY[1*], **İsmihan BAYRAMOĞLU**[2]
[1]Department of Statistics, Faculty of Art and Science, Yıldız Technical University, İstanbul, Turkey
[2]Department of Mathematics, Faculty of Art and Science, İzmir University of Economics, İzmir, Turkey

Received: 14.01.2019 • **Accepted/Published Online:** 26.03.2019 • **Final Version:** 29.05.2019

Abstract: The theory of record values has been extensively studied in the statistical literature. However, there are not many papers devoted to the theory of records for bivariate and multivariate random sequences. This paper presents the marginal record values and record times in extended sequence of bivariate random vectors. The joint distributions of some upper marginal records are derived. Some results on joint probability mass function of upper record time vectors and distribution function of upper record value vectors are given via copula functions. Moreover, the numerical and graphical applications of considered upper records using flood data and prediction of rainfall variables such as intensity, depth, and duration are provided.

Key words: Bivariate records, copula, joint distribution function, flood data, record values, record times

1. Introduction

Interest in the theory of records has increased rapidly since the appearance of the first pioneering paper by Chandler [11]. Record values and record times have wide applications in many areas, including sport events, hydrology, extreme weather occurrences, biology, engineering, and economics. There are many papers devoted to distributional properties of record times and record values for general classes as well as for specific distributions. See, e.g., [2, 3, 19, 20], among others. Some papers study the characterization of distributions through the properties of record values. For example, the characterization properties of the exponential distribution were studied in [1]. Some characterizations of the geometric distribution were presented in [5]. There are also fundamental books that made tremendous contributions to this area. In [4], Ahsanullah gave a comprehensive review of the field of records. A detailed description of the theory and applications of records can be found in [6]. The mathematical foundation of the theory of records was given in [21]. Although the theory of records has been well developed, only a few papers in the statistical literature deal with bivariate and multivariate records. Some of these works introduce the definition of records according to different ordering principles of multivariate random vectors. Based on partial ordering, Goldie and Resnick [16] considered a sequence of independent and identical (iid) $\mathbb{R}^2$-valued random vectors $X_n = (X_n^1, X_n^2), n = 1, 2, ...$ and called X_n a record if there is a record in both coordinates. Similar concepts of multivariate records in partially ordered sets and other notions were discussed in [14, 15, 17]. In recent years, [7] considered records of bivariate sequences by using conditional ordering and has gave the definition of multivariate records according to component-wise ordering.

In this paper, we consider a bivariate random sequence of independent random vectors $(X_1, Y_1), (X_2, Y_2), ...$

*Correspondence: kemalbay@yildiz.edu.tr
2010 *AMS Mathematics Subject Classification:* 62E15, 60E05

with absolutely continuous joint distribution function (cdf) F. Taking into account the records of the sequences $\{X_k\}_{k\geq 1}$ and $\{Y_k\}_{k\geq 1}$, we study the joint probability mass function (pmf) of record times and joint probability density function (pdf) of record values from these two sequences. We consider an extended sequence $\{(X_k, Y_l)\}_{k\geq 1, l\geq 1}$ and define the bivariate (n,m)th record time vector as $(T(n), S(m))$ and the corresponding (n,m)th record value vector as $(X_{T(n)}, Y_{S(m)})$, respectively. Finding a joint distribution of the above-mentioned record time vector and record value vector for any $n > 2$ and $m > 2$ appears to be a challenging problem. In the present paper, the joint pmf of $(T(2), S(2))$ and the joint pdf of $(X_{T(2)}, Y_{S(2)})$ have been derived, and some of the results are expressed in terms of copula functions of bivariate random vectors. Some examples for special distributions including the independence case are provided.

Record values and record times arise naturally in both theoretical and practical areas of probability and statistics. On the theoretical side, for example, knowledge of exact distribution functions of the record value sequence is sufficient to characterize the common distribution function of the underlying observations. Moreover, various statistical inference procedures such as hypothesis testing and point or interval estimation and prediction can be developed based on observed record sequences. On the practical side, record values and record times have attracted great attention when analyzing extreme weather and climate events or studying sports, traffic, medicine, and financial events, among others. The prediction of future record values and record times on the basis of the past records is of interest in many areas, especially in meteorology and hydrology (see, e.g., [8, 9]). For example, in floodplain management, the joint distribution of rainfall variables (e.g., intensity, depth, duration) is important for reducing flood risks and protecting healthy ecosystems. Assuming the extended bivariate sequence of dependent random variables $\{(X_k, Y_l)\}_{k\geq 1, l\geq 1}$, where X_i denotes the rainfall intensity and Y_i denotes the rainfall depth in the ith year, the record data in X and Y are an example of bivariate records. If the record of rainfall intensity for a given year is known, it will be naturally relevant in predicting the record of rainfall depth. Under minimum mean squared error, the best unbiased predictor for $T(n)$ given $S(m), m < n$ is the conditional expectation $E\{T(n) \mid S(m)\}$ and similarly $E\{X_{T(n)} \mid Y_{S(m)}\}$ is the best predictor of $X_{T(n)}$ given $Y_{S(m)}$. Therefore, while predicting the value or time of the next record, we need the joint pmf of record times and joint cdf of record values, which are the main subjects of this paper.

2. Records of extended bivariate random sequence

Let us briefly describe the univariate record times and values for better understanding the bivariate records we are dealing in this paper. Suppose that $\{X_k\}_{k\geq 1}$ is a sequence of iid random variables from an absolutely continuous distribution. The classical univariate upper record time $T(n)$ and upper record value $X_{T(n)}$ of the sequence $\{X_k\}_{k\geq 1}$ are defined as follows:

$$\begin{aligned} &T(1) = 1;\ X_{T(1)} = X_1, \\ &T(n) = \min\{i \mid i > T(n-1), X_i > X_{T(n-1)}\}, \forall n \geq 2. \end{aligned} \tag{2.1}$$

Many distributional properties of record values in the sequence of iid absolutely continuous random variables $X_1, X_2, ...$ with cdf $F_X(x)$ and pdf $f(x)$ have been expressed in terms of the function $R(x) = -\ln[1 - F_X(x)]$

(see [4]). The distribution function of $X_{T(n)}$, $n \geq 2$, denoted by $F_n(x)$ is

$$\begin{aligned} F_n(x) &= P\{X_{T(n)} \leq x\} \\ &= \frac{1}{\Gamma(n)} \int_{-\infty}^{R(x)} u^{n-1} e^{-u} du \\ &= 1 - e^{-R(x)} \sum_{j=0}^{n-1} \frac{[R(x)]^j}{j!}, \end{aligned} \tag{2.2}$$

and the pdf of $X_{T(n)}$ denoted by $f(x_n)$ is

$$f_n(x) = \frac{[R(x)]^{n-1}}{\Gamma(n)} f(x), \tag{2.3}$$

where $\Gamma(n)$ is a gamma function and $\Gamma(n) = (n-1)!$.

Four definitions of bivariate records were described in [6]. Let $\{(X_k, Y_k)\}_{k\geq 1}$ be a sequence of bivariate random vectors. According to Definition 8.1.4 in [6, p. 266], for example, the observation (X_k, Y_k) is a record if X_k precedes all X_is for $i < k$ and Y_k precedes all Y_is for $i < k$. In this paper, we are interested in marginal records of each of the sequences $\{X_k\}_{k\geq 1}$ and $\{Y_k\}_{k\geq 1}$ and we investigate the joint distributions of marginal record times and record values of these sequences. More precisely, let (X, Y) be a random vector with absolutely continuous cdf $F(x, y)$ and pdf $f(x, y)$, where $F_X(x)$, $F_Y(y)$ are corresponding marginal cdfs and $f_X(x)$, $f_Y(y)$ are corresponding marginal pdfs. Consider a sequence of random vectors $\{(X_k, Y_k)\}_{k\geq 2}$ where (X_k, Y_k) are independent copies of (X, Y). Alternatively, we can define bivariate records in an extended sequence constructed from the sequences $\{X_k\}_{k\geq 1}$ and $\{Y_k\}_{k\geq 1}$ as follows: let $\{(X_k, Y_l)\}_{k\geq 1, l\geq 1}$ be an extended sequence of random variables constructed from copies of X and Y. We say that (X_k, Y_l) is an upper record value vector if (X_k, Y_l) is an upper record value with respect to each X and Y simultaneously. In other words, the observation (X_k, Y_l) is a record if X_k precedes all X_is for $i < k$ and Y_l precedes all Y_is for $i < l$. The upper record time vector (k, l) consisting of indices at which upper record values occur can be realized in the following cases. The first coordinate of upper record value X_k may occur before Y_l, which means that $k < l$; the second coordinate Y_l may occur before X_k, i.e. $l < k$; and they may occur simultaneously, i.e. $k = l$. To formalize this notion, we give the following definition.

Definition 2.1 *Assume that $\{(X_k, Y_k)\}_{k\geq 1}$ is a sequence of iid bivariate random vectors with continuous cdf $F(x, y)$. Consider an extended bivariate random sequence $\{(X_k, Y_l)\}_{k\geq 1, l\geq 1}$, where (X_k, Y_k) are independent copies of (X, Y). The (n, m)th upper record time vector of the corresponding sequence denoted by $(T(n), S(m))$ is given as follows:*

$$(T(n), S(m)) = (\min\{i \mid i > T(n-1), X_i > X_{T(n-1)}\}, \min\{j \mid j > S(m-1), Y_j > Y_{S(m-1)}\}), \forall n, m \geq 2. \tag{2.4}$$

Then $(X_{T(n)}, Y_{S(m)})$ is called the (n, m)th upper record value vector of the extended sequence if

$$(X_{T(n)}, Y_{S(m)}) > (\max(X_1, X_2, ..., X_{T(n-1)}), \max(Y_1, Y_2, ..., Y_{S(m-1)})). \tag{2.5}$$

By assumption, the first upper record time vector of the extended sequence is $(T(1), S(1)) = (1, 1)$ and the first upper record value vector is $(X_{T(1)}, Y_{S(1)}) = (X_1, Y_1)$. All possible orderings between the coordinates of upper record time vector $(T(n), S(m))$ could be as follows: $T(n) < S(m)$ or $T(n) > S(m)$ or $T(n) = S(m)$.

Note that the (k,k)th record value vector of the extended sequence $\{(X_k, Y_l)\}_{k\geq 1, l\geq 1}$ given in Definition 2.1 is actually the kth record value vector of the sequence $\{(X_k, Y_k)\}_{k\geq 1}$ according to Definition 8.1.4 in [6, p.266].

To illustrate our definition, we give a simple example. The average temperatures in Switzerland in July and August during 1985–1993 are listed in the following table. From Table 1, according to nine observations, for example, the $(3,2)$nd upper record time vector is $(6,5)$ and the corresponding record value vector is $(20.3, 19.2)$.

Table 1. The average temperatures in Switzerland during 1985–1993 [6].

	1985	1986	1987	1988	1989	1990	1991	1992	1993
July (X)	19.2	20.2	19.3	19.0	18.8	20.3	19.7	20.7	19.6
August (Y)	18.6	18.4	18.1	18.6	19.2	19.1	20.2	21.1	21.5

2.1. Joint probability mass function of marginal records

Using the definition of the (n,m)th record time vector and copula structure of bivariate distributions, one can derive the joint pmf of the (n,m)th record time vector. The following theorem provides the pmf of the upper record time vector for $n = m = 2$ and states that this pmf can be expressed in terms of the copula; it does not depend on marginal distributions of X and Y.

Theorem 2.2 *Let $\{(X_k, Y_k)\}_{k\geq 1}$ be a sequence of iid bivariate random vectors with absolutely continuous cdf $F(x,y) = C(F_X(x), F_Y(y))$ where $C(u,v)$, $[u,v] \in [0,1]^2$, is the connecting copula. Hence, the joint pdf is $f(x,y) = c(F_X(x),\ F_Y(y))f_X(x)f_Y(y)$, where $c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}$ is the copula density function. Assume that $F_X(x)$ and $F_Y(y)$ are strictly increasing marginal cdfs. Consider an extended sequence $\{(X_k, Y_l)\}_{k\geq 1, l\geq 1}$ and the $(2,2)$nd upper record time vector $(T(2), S(2))$. The joint pmf of $(T(2), S(2))$ is*

$$P\{T(2) = i, S(2) = j\} = \begin{cases} \int_0^1\int_0^1 [C(u,v)]^{i-2}[v - C(u,v)]v^{j-i-1}[1-v]c(u,v)dudv, & i < j \\ \int_0^1\int_0^1 [C(u,v)]^{j-2}[u - C(u,v)]u^{i-j-1}[1-u]c(u,v)dudv, & i > j \\ \int_0^1\int_0^1 [C(u,v)]^{i-2}[1 - u - v + C(u,v)]c(u,v)dudv, & i = j \end{cases}.$$

Proof According to the Definition 2.1, one can write

$$\begin{aligned} P\{T(2) = i, S(2) = j\} &= P\{X_1 > X_2, ..., X_1 > X_{i-1}, X_1 < X_i, Y_1 > Y_2, ..., Y_1 > Y_{j-1}, Y_1 < Y_j\} \\ &= P\{F_X(X_1) > F_X(X_2), ..., F_X(X_1) > F_X(X_{i-1}), F_X(X_1) < F_X(X_i), F_Y(Y_1) > F_Y(Y_2), \\ &\quad ..., F_Y(Y_1) > F_Y(Y_{j-1}), F_Y(Y_1) < F_Y(Y_j)\} \\ &= P\{U_1 > U_2, ..., U_1 > U_{i-1}, U_1 < U_i, V_1 > V_2, ..., V_1 > V_{j-1}, V_1 < V_j\}, \end{aligned}$$

where $F_X(X) = U$, and $F_Y(Y) = V$ and $P\{U \le u, V \le v\} = C(u,v)$. According to orderings between i and j indices at which the second upper record values occur, we consider cases $i < j$, $i > j$, and $i = j$.

If $i < j$, then

$$P\{T(2) = i, S(2) = j\}$$
$$= P\{U_1 > U_2, ..., U_1 > U_{i-1}, U_1 < U_i, V_1 > V_2, ..., V_1 > V_{i-1}, V_1 > V_i, ..., V_1 > V_{j-1}, V_1 < V_j\}$$
$$= \int_0^1 \int_0^1 P\{U_1 > U_2, ..., U_1 > U_{i-1}, U_1 < U_i, V_1 > V_2, ..., V_1 > V_{i-1},$$
$$V_1 > V_i, ..., V_1 > V_{j-1}, V_1 < V_j \mid U_1 = u, V_1 = v\} dC(u,v)$$
$$= \int_0^1 \int_0^1 P\{U_2 < u, V_2 < v, ..., U_{i-1} < u, V_{i-1} < v; U_i > u, V_i < v; V_{i+1} < v, ..., V_{j-1} < v; V_j > v\} dC(u,v)$$
$$= \int_0^1 \int_0^1 [C(u,v)]^{i-2} [v - C(u,v)] v^{j-i-1} [1-v] c(u,v) du dv.$$

If $i > j$, then

$$P\{T(2) = i, S(2) = j\}$$
$$= P\{U_1 > U_2, ..., U_1 > U_{j-1}, U_1 > U_j, ..., U_1 > U_{i-1}, U_1 < U_i, V_1 > V_2, ..., V_1 > V_{j-1}, V_1 < V_j\}$$
$$= \int_0^1 \int_0^1 P\{U_1 > U_2, ..., U_1 > U_{j-1}, U_1 > U_j, ..., U_1 > U_{i-1}, U_1 < U_i,$$
$$V_1 > V_2, ..., V_1 > V_{j-1}, V_1 < V_j \mid U_1 = u, V_1 = v\} dC(u,v)$$
$$= \int_0^1 \int_0^1 P\{U_2 < u, V_2 < v, ..., U_{j-1} < u, V_{j-1} < v; U_j < u, V_j > v; U_{j+1} < u, ..., U_{i-1} < u; U_i > u\} dC(u,v)$$
$$= \int_0^1 \int_0^1 [C(u,v)]^{j-2} [u - C(u,v)] \; u^{i-j-1} [1-u] c(u,v) du dv.$$

If $i = j$, then

$$P\{T(2) = i, S(2) = j\}$$
$$= P\{U_1 > U_2, ..., U_1 > U_{i-1}, U_1 < U_i, V_1 > V_2, ..., V_1 > V_{i-1}, V_1 < V_i\}$$
$$= \int_0^1 \int_0^1 P\{U_1 > U_2, ..., U_1 > U_{i-1}, U_1 < U_i, V_1 > V_2, ..., V_1 > V_{i-1}, V_1 < V_i \mid U_1 = u, V_1 = v\} dC(u,v)$$

$$= \int_0^1 \int_0^1 P\{U_2 < u, V_2 < v, ..., U_{i-1} < u, V_{i-1} < v; U_i > u,\ V_i > v\} dC(u,v)$$

$$= \int_0^1 \int_0^1 [C(u,v)]^{i-2}[1-u-v+C(u,v)]c(u,v)dudv.$$

□

Since $T(n)$ and $S(m)$ are nth and mth record times of $\{X_k\}_{k\geq 1}$ and $\{Y_k\}_{k\geq 1}$, respectively, it is interesting to calculate the probability of event $A_{n,m} = \{n$th record of X sequence comes before the mth record of Y sequence, i.e. $T(n) < S(m)\}$, for any $n \geq 2, m \geq 2$. In particular, the probability $P\{T(2) < S(2)\}$ is given below in Proposition 2.3.

Proposition 2.3

$$P\{T(2) < S(2)\} = \sum_{i=2}^{\infty} \sum_{j=i+1}^{\infty} \int_0^1 \int_0^1 [C(u,v)]^{i-2}[v - C(u,v)]v^{j-i-1}[1-v]c(u,v)dudv$$

$$= \int_0^1 \int_0^1 \frac{[v - C(u,v)]}{[1 - C(u,v)]} c(u,v)dudv.$$

Similarly,

$$P\{T(2) > S(2)\} = \int_0^1 \int_0^1 \frac{[u - C(u,v)]}{[1 - C(u,v)]} c(u,v)dudv,$$

and

$$P\{T(2) = S(2)\} = \int_0^1 \int_0^1 \frac{[1 - u - v + C(u,v)]}{[1 - C(u,v)]} c(u,v)dudv.$$

Note that the joint pmf of $(T(2), S(2))$ in the case of $i < j$ and $i > j$ holds the same for exchangeable random variables, i.e. symmetric copulas satisfying $C(u,v) = C(v,u), \forall u,v \in [0,1]^2$.

In floodplain management, the probability of second upper record times gains importance to calculate the risk of extreme flooding events. The influence of the association parameter on the probability functions given in Proposition 2.3 is studied in the following example by considering flood data from the Gumbel–Hougaard copula.

Example 2.4 *Consider $C(u,v) = \exp(-[(-\ln u)^\alpha + (-\ln v)^\alpha]^{1/\alpha}), \alpha \geq 1$, the Gumbel–Hougaard copula that corresponds to type B bivariate extreme-value distribution $F(x,y) = \exp(-[(-\ln F_X(x))^m + (-\ln F_Y(y))^m]^{1/m})$, known as the logistic model. In flood frequency analysis, one of the most fundamental assumptions is that either flood variables such as rainfall intensity, depth, duration, or stations are independent or each have the same marginal probability distribution, usually assumed to be normal. However, in reality, floods for an upstream and a downstream stations are usually highly dependent and do not follow normal distribution, which presents evidence of extreme value distribution. The Gumbel–Hougaard copula is the only one known to be both Archimedean* *and an extreme value copula and it does not assume the variables to be independent or have the same type of marginal distributions. It contains the parameter* α*, which shows the degree of association between the extremes. To show the influence of parameter* α *on the pmf of* $(2,2)$*nd upper record time vector, we plot the probabilities given in Proposition 2.3 in the same graph versus different values of parameter* α *in Figure 1.*

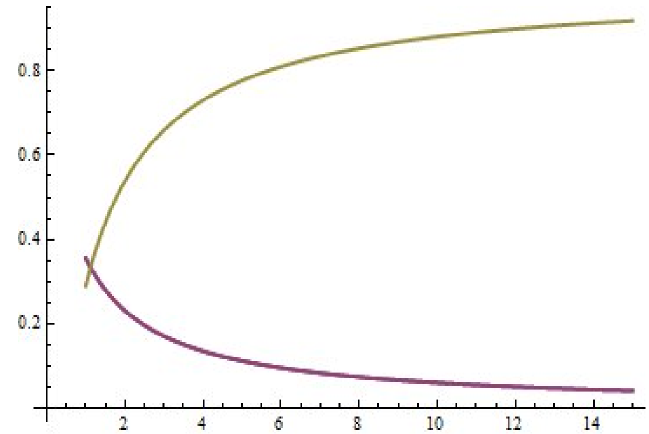

Figure 1. The graph of probabilities given in Proposition 2.3 for Gumbel–Hougaard copula versus several values of α.

From Figure 1, it can be said that if parameter α *increases, then the probability of the second upper record values occurring simultaneously will increase (the yellow curve in the graph) and the probability of the second upper record value from the* X *sequence coming before or after the second upper record value from the* Y *sequence will decrease (the purple curve in the graph shows the same for* $P\{T(2) < S(2)\}$ *and* $P\{T(2) > S(2)\}$ *since the Gumbel–Hougaard copula is symmetric). Gumbel and Mustafi [18] fitted the type B bivariate extreme value distribution to data on the floods of the Fox River upstream at Berlin and downstream at Wrightstown from the year* 1918 *to* 1950 *and the parameter* α *was estimated by the modified method of moments as* $\widehat{\alpha} = 1.780$*. Considering the same data, we compare the results of Proposition 2.3 for the Gumbel–Hougaard copula for fitted* α *and* $\alpha = 1$*, i.e. the independence case.*

Table 2. Pmf values of the $(2,2)$nd upper record time vector for data given in [18] with fitted α and independence case of α.

Prob.	$\widehat{\alpha} = 1.780$	$\alpha = 1$
$P\{T(2) < S(2)\}$	0.250	0.355
$P\{T(2) > S(2)\}$	0.250	0.355
$P\{T(2) = S(2)\}$	0.499	0.290

According to Table 2, for example, the risk of the simultaneous occurrence of the second upper record values belonging to the data of the floods of the Fox River upstream and floods of the Fox River downstream will increase evidently from 0.290 *to* 0.499 *if one takes into account the dependency.*

For multivariate analysis, the dependence concept between random variables is important because the independence assumption is rarely available in practice. Among the bivariate dependence concepts, the notion of positive quadrant dependent (PQD) states that the random variables are more likely to be large (or small) simultaneously compared to the independence case. The Farlie–Gumbel–Morgenstern (FGM) family of copulas is one of the important classes of copulas possessing the PQD property. Therefore, in modeling bivariate data, many researchers prefer the FGM family of distributions, which allows analysis of the dependence properties of bivariate records and record times. Below, in Example 2.5, in order to illustrate the result of Theorem 2.2 in the case of dependency between X and Y, we consider the FGM copula that involves the association parameter α.

Example 2.5 *Let us consider FGM copula* $C(u,v) = uv[1+\alpha(1-u)(1-v)]$, $(u,v) \in [0,1]^2$, $-1 \le \alpha \le 1$. *It is known that the FGM copula has the positive quadrant dependent property for* $0 \le \alpha \le 1$. *Furthermore, U and V have maximal negative dependence for* $\alpha = -1$ *and maximal positive dependence for* $\alpha = 1$. *While* α *changes from* -1 *to* 1, *the positive dependence increases and negative dependence decreases for the underlying FGM copula. Below, we provide the graph of the values of* $P\{T(2) = i, S(2) = j\}$ *for different* i *and* j *versus association parameter* α.

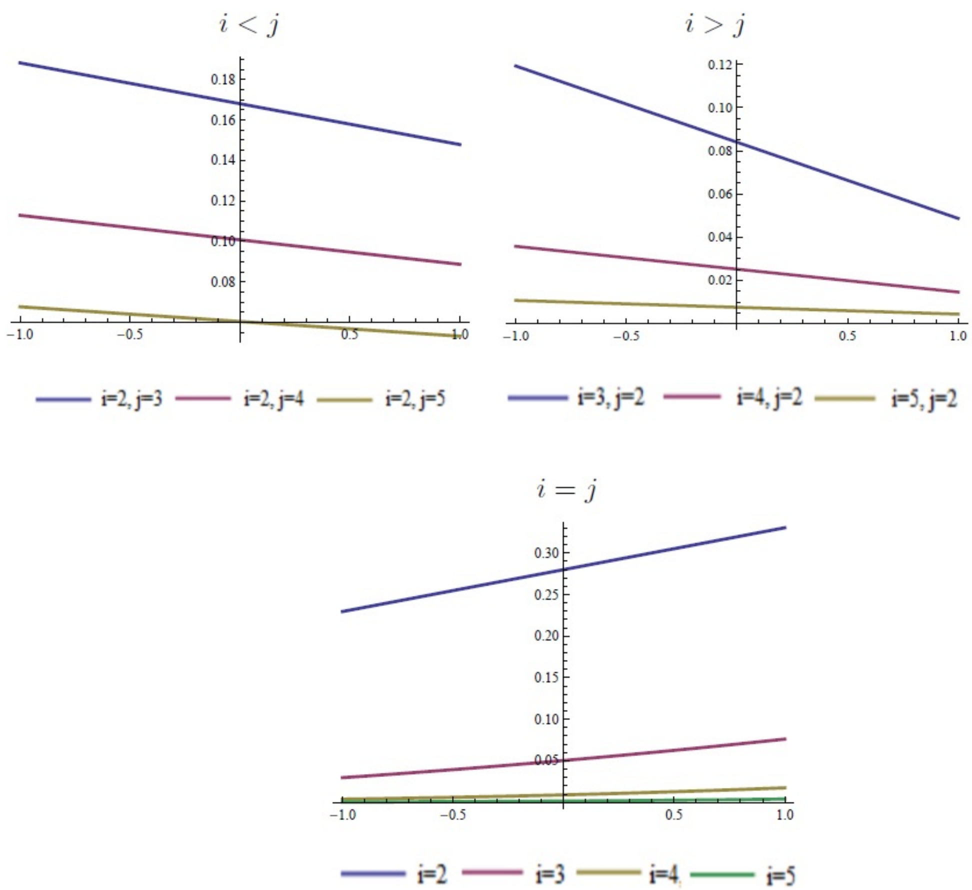

Figure 2. The graph of pmf values of $P\{T(2) = i, S(2) = j\}$ of FGM copula for different values of i and j plotted versus α.

From Figure 2, we can say that when α *increases from* -1 *to* 1, *then the probability* $P\{T(2) = i, S(2) = j\}$ *decreases for both cases* $i < j$ *and* $i > j$ *and increases for* $i = j$. *Therefore, the increase in the positive dependence leads to a decrease in the pmf of the* $(2,2)$*nd upper record time vector when* $i < j$ *or* $i > j$ *for this type of copula. However, if the marginal record values from the bivariate sample occur at the same time, i.e.* $i = j$, *then the pmf increases when* α *increases.*

The following theorem provides the joint pmf of $T(2), S(2)$, and $T(3)$.

Theorem 2.6 *Assume that* $\{(X_k, Y_k)\}_{k\geq 1}$ *is a sequence of iid bivariate random vectors with continuous cdf* $F(x,y) = C(F_X(x), F_Y(y))$ *and* $F_X(x)$, $F_Y(y)$ *are the corresponding strictly increasing marginal cdf of the* X *and* Y *sample, respectively. Consider the trivariate random vector* $(T(2), S(2), T(3))$ *in the extended sequence* $\{(X_k, Y_l)\}_{k\geq 1, l\geq 1}$. *Then the joint pmf of upper record time vector* $(T(2), S(2), T(3))$ *is*

$$P\{T(2)=i, S(2)=j, T(3)=k\}$$

$$= \begin{cases} \int_0^1\int_0^1\int_0^z [C(u,v)]^{i-2} v[C(z,v)]^{j-i-1}[z-C(z,v))]z^{k-j-1}[1-z]c(u,v)dudvdz, & i<j<k \\ \int_0^1\int_0^1\int_0^z [C(u,v)]^{i-2}[C(z,v)]^{k-i-1}[v-C(z,v))]v^{j-k}[1-v]c(u,v)dudvdz, & i<k<j \\ \int_0^1\int_0^1\int_0^z [C(u,v)]^{i-2} v[C(z,v)]^{k-i-1}[1-z-v+C(z,v)]c(u,v)dudvdz, & i<j=k \\ \int_0^1\int_0^1\int_0^z [C(u,v)]^{j-2}[u-C(u,v)]u^{i-j-1}z^{k-i-1}[1-z]c(u,v)dudvdz, & j<i<k \\ \int_0^1\int_0^z [C(u,v)]^{i-2}[1-v]z^{k-i-1}[1-z]c(u,v)dudvdz, & i=j<k \end{cases}.$$

Proof

$$\begin{aligned}
&P\{T(2)=i, S(2)=j, T(3)=k\} \\
&= P\{X_1>X_2,...,X_1>X_{i-1}, X_1<X_i, Y_1>Y_2,...,Y_1>Y_{j-1}, Y_1<Y_j, X_i>X_{i+1},...,X_i>X_{k-1}, X_i<X_k\} \\
&= P\{F_X(X_1)>F_X(X_2),...,F_X(X_1)>F_X(X_{i-1}), F_X(X_1)<F_X(X_i), F_Y(Y_1)>F_Y(Y_2),..., \\
&F_Y(Y_1)>F_Y(Y_{j-1}), F_Y(Y_1)<F_Y(Y_j), F_X(X_i)>F_X(X_{i+1}),...,F_X(X_i)>F_X(X_{k-1}), F_X(X_i)<F_X(X_k)\} \\
&= P\{U_1>U_2,...,U_1>U_{i-1}, U_1<U_i, V_1>V_2,...,, V_1>V_{j-1}, V_1<V_j, U_i>U_{i+1},...,U_i>U_{k-1}, U_i<U_k\},
\end{aligned}$$

where $F_X(X)=U$, and $F_Y(Y)=V$ and $P\{U\leq u, V\leq v\}=C(u,v)$. By conditioning upon upper record times $U_1=u, V_1=v$, and $U_i=z$, we get

$$P\{T(2)=i,S(2)=j,T(3)=k\}$$

$$=\int_0^1\int_0^1\int_0^1 P\{U_1>U_2,...,U_1>U_{i-1},U_1<U_i,V_1>V_2,...,,V_1>V_{j-1},V_1<V_j,$$

$$U_i>U_{i+1},...,U_i>U_{k-1},U_i<U_k\mid U_1=u,U_2=v,U_i=z\}c(u,v)dudvdz$$

$$=\int_0^1\int_0^1\int_0^z P\{U_2<u,...,U_{i-1}<u,u<z,V_2<v,...,V_{j-1}<v,V_j>v,U_{i+1}<z,...,U_{k-1}<z,U_k>z\}$$

$$\times c(u,v)dudvdz.$$

Let us consider the sequence of upper record time vector $\{(T(n),S(m));\ n,m\geq 2\}$ in extended sequence $\{(X_k,Y_l)\}_{k\geq 1,l\geq 1}$. Specifically, for the trivariate random vector $(T(2),S(2),T(3))$, all possible orderings among the indices i,j, and k can be such that: *case* 1: $i<j<k$; *case* 2: $i<k<j$; *case* 3: $i<j=k$; *case* 4: $j<i<k$; *case* 5: $i=j<k$.

For *case* 1: $i<j<k$,

$$P\{T(2)=i,S(2)=j,T(3)=k\}$$

$$=\int_0^1\int_0^1\int_0^z P\{U_2<u,V_2<v,...,U_{i-1}<u,V_{i-1}<v;u<z,V_i<v;U_{i+1}<z,V_{i+1}<v,...,U_{j-1}<z,V_{j-1}<v;$$

$$U_j<z,V_j>v;U_{j+1}<z,...,U_{k-1}<z;U_k>z\}c(u,v)dudvdz$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{z}[C(u,v)]^{i-2}v[C(z,v)]^{j-i-1}[z-C(z,v)]z^{k-j-1}[1-z]c(u,v)dudvdz.$$

For *case* 2: $i<k<j$,

$$P\{T(2)=i,S(2)=j,T(3)=k\}$$

$$=\int_0^1\int_0^1\int_0^z P\{U_2<u,V_2<v,...,U_{i-1}<u,V_{i-1}<v;u<z,V_i<v;$$

$$U_{i+1}<z,V_{i+1}<v,...,U_{k-1}<z,V_{k-1}<v;U_k>z,V_k<v;V_{k+1}<v,...,V_{j-1}<v;V_j>v\}c(u,v)dudvdz$$

$$=\int_0^1\int_0^1\int_0^z[C(u,v)]^{i-2}[C(z,v)]^{k-i-1}[v-C(z,v)]v^{j-k}[1-v]c(u,v)dudvdz.$$

For *case* 3: $i < j = k$,

$$P\{T(2) = i, S(2) = j, T(3) = k\}$$

$$= \int_0^1 \int_0^1 \int_0^z P\{U_2 < u, V_2 < v, ..., U_{i-1} < u, V_{i-1} < v; u < z, V_i < y; U_{i+1} < z, V_{i+1} < v, ..., U_{k-1} < z, V_{k-1} < v;$$

$$U_k > z, V_k > v\}c(u,v)dudvdz$$

$$= \int_0^1 \int_0^1 \int_0^z [C(u,v)]^{i-2} v [C(z,v)]^{k-i-1} [1 - z - v + C(z,v)] c(u,v) dudvdz.$$

For *case* 4: $j < i < k$,

$$P\{T(2) = i, S(2) = j, T(3) = k\}$$

$$= \int_0^1 \int_0^1 \int_0^z P\{U_2 < u, V_2 < v, ..., U_{j-1} < u, V_{j-1} < v; U_j < u, V_j > v; U_{j+1} < u, ..., U_{i-1} < u; U_i > u; U_{i+1} < z, ...,$$

$$U_{k-1} < z; U_k > z\}c(u,v)dudvdz$$

$$= \int_0^1 \int_0^1 \int_0^z [C(u,v)]^{j-2} [u - C(u,v)] u^{i-j-1} z^{k-i-1} [1 - z] c(u,v) dudvdz.$$

For *case* 5: $i = j < k$,

$$P\{T(2) = i, S(2) = j, T(3) = k\}$$

$$= \int_0^1 \int_0^1 \int_0^z P\{U_2 < u, V_2 < v, ..., U_{i-1} < u, V_{i-1} < v; u < z, V_i > v; U_{i+1} < z, ..., U_{k-1} < z; U_k > z\}c(u,v)dudvdz$$

$$= \int_0^1 \int_0^1 \int_0^z [C(u,v)]^{i-2} [1 - v] z^{k-i-1} [1 - z] c(u,v) dudvdz.$$

□

Now we are interested in the probability of the event $\{T(2)$ is less than $S(2)$ and $S(2)$ is less than $T(3)\}$, i.e. $P\{T(2) < S(2) < T(3)\}$. We compute the following probabilities according to all possible orderings among $T(2)$, $S(2)$, and $T(3)$ in Proposition 2.7.

Proposition 2.7

$$P\{T(2) < S(2) < T(3)\}$$

$$= \int_0^1 \int_0^1 \int_0^z \frac{v[z - C(z,v)]}{[1 - C(u,v)][1 - C(z,v)]} c(u,v) dudvdz.$$

$$P\{T(2) < T(3) < S(2)\}$$
$$= \int_0^1 \int_0^1 \int_0^z \frac{v[v - C(z,v)]}{[1 - C(u,v)][1 - C(z,v)]} c(u,v) du dv dz.$$

$$P\{T(2) < S(2) = T(3)\}$$
$$= \int_0^1 \int_0^1 \int_0^z \frac{v[1 - z - v + C(z,v)]}{[1 - C(u,v)][1 - C(z,v)]} c(u,v) du dv dz.$$

$$P\{S(2) < T(2) < T(3)\} \equiv P\{S(2) < T(2)\}$$
$$= \int_0^1 \int_0^1 \int_0^z \frac{[u - C(u,v)]}{[1 - u][1 - C(u,v)]} c(u,v) du dv dz.$$

$$P\{T(2) = S(2) < T(3)\} \equiv P\{T(2) = S(2)\}$$
$$= \int_0^1 \int_0^1 \int_0^z \frac{[1 - v]}{[1 - C(u,v)]} c(u,v) du dv dz.$$

Corollary 2.8 *As a result of Theorem 2.6, we can obtain the joint pmf of* $(T(3), S(2))$ *as follows:*

$$P\{T(3) = k, S(2) = j\}$$
$$= \begin{cases} \int_0^1 \int_0^1 \int_0^z ([C(z,v)]^k [C(u,v)]^2 - [C(z,v)]^2 [C(u,v)]^k)[v - C(z,v)] \\ \times [1 - v] v^{j-k} \frac{1}{[C(z,v)]^2 [C(z,v) - C(u,v)][C(u,v)]^2} c(u,v) du dv dz, & k < j \\ \int_0^1 \int_0^1 \int_0^z [C(z,v)]^{j-k-2} ([C(z,v)]^k [C(u,v)]^2 - [C(z,v)]^2 [C(u,v)]^k) \\ \times [z - C(z,v)][1 - z] z^{k-j-1} \frac{1}{[C(z,v) - C(u,v)][C(u,v)]^2} c(u,v) du dv dz, & k > j \\ \int_0^1 \int_0^1 \int_0^z ([C(z,v)]^k [C(u,v)]^2 - [C(z,v)]^2 v[1 - z - v + C(z,v)] \\ \times \frac{1}{[C(z,v)]^2 [C(z,v) - C(u,v)][C(u,v)]^2} c(u,v) du dv dz, & k = j \end{cases}.$$

In Proposition 2.9, we obtain the probabilities that $T(3)$ is less than $S(2)$, $T(3)$ is greater than $S(2)$, and $T(3)$ is equal to $S(2)$, respectively, as follows.

Proposition 2.9 *It is clear that*

$$P\{T(3) < S(2)\}$$
$$= \int_0^1\int_0^1\int_0^z \frac{v[v - C(z,v)]}{[1 - C(u,v)][1 - C(z,v)]} c(u,v)dudvdz.$$

$$P\{T(3) > S(2)\}$$
$$= \int_0^1\int_0^1\int_0^z \frac{v[z - C(z,v)][1 - zC(u,v)]}{[1 - C(u,v)][1 - C(z,v)][C(z,v)] - zC(u,v)]} c(u,v)dudvdz.$$

$$P\{T(3) = S(2)\}$$
$$= \int_0^1\int_0^1\int_0^z \frac{v[1 - z - v + C(z,v)]}{[1 - C(u,v)][1 - C(z,v)]} c(u,v)dudvdz.$$

In the next theorem, we obtain the joint pmf of $(T(2), S(2), S(3))$.

Theorem 2.10 *Assume that* $\{(X_k, Y_k)\}_{k\geq 1}$ *is a sequence of iid bivariate random vectors with continuous cdf* $F(x,y) = C(F_X(x), F_Y(y))$ *and* $F_X(x)$*, and* $F_Y(y)$ *is the corresponding strictly increasing marginal cdf of the* X *and* Y *sample, respectively. Consider the trivariate random vector* $(T(2), S(2), S(3))$ *in the extended sequence* $\{(X_k, Y_l)\}_{k\geq 1, l\geq 1}$*. Then the joint pmf of upper record time vector* $(T(2), S(2), S(3))$ *is*

$$P\{T(2) = i, S(2) = j, S(3) = l\}$$
$$= \begin{cases} \int_0^1\int_0^w\int_0^1 [C(u,v)]^{j-2} u[C(u,w)]^{l-j-1}[u - C(u,w)]u^{i-l-1}[1-u]c(u,v)dudvdw, & j < l < i \\ \int_0^1\int_0^w\int_0^1 [C(u,v)]^{j-2} u[C(u,w)]^{i-j-1}[w - C(u,w)]w^{l-i-1}[1-w]c(u,v)dudvdw, & j < i < l \\ \int_0^1\int_0^w\int_0^1 [C(u,v)]^{j-2}[1 - u - w + C(u,w)]u[C(u,w)]^{i-j-1}c(u,v)dudvdw, & j < i = l \\ \int_0^1\int_0^w\int_0^1 [C(u,v)]^{i-2}[v - C(u,v)]v^{j-i-1}w^{l-j-1}[1-w]c(u,v)dudvdw, & i < j < l \\ \int_0^1\int_0^w\int_0^1 [C(u,v)]^{j-2}[1-u]w^{l-j-1}[1-w]c(u,v)dudvdw, & i = j < l \end{cases}.$$

Proof The proof is similar to the proof of Theorem 2.6. □

To illustrate the results of Theorem 2.6, we consider a numerical example in hydrology. Multivariate hydrological events such as floods, rainfalls, and storms that can each be characterized with a few correlated variables and the dependence between these variables have been classically modeled by bivariate gamma, log-normal, and extreme value distributions as well as by the copula approach (see, e.g., [10, 12, 13]).

In the next example, the effect of Kendall's τ rank correlation coefficient on pmf values of $(T(3),\ S(2))$ given in Proposition 2.9 has been investigated independently from the choice of marginal distributions through the copula approach. The FGM copula is chosen for its simplicity and flexibility since it is adequate for the data that exhibit moderate dependence and could capture both negative and positive dependences between the marginals.

Example 2.11 *The flood frequency analysis is often represented by the variables of flood peak and flood volume, which are usually positively correlated. Suppose that X_i denotes the flood peak and Y_i denotes the flood volume in the ith year. In this example, for the joint modeling of record times data in X and Y, we consider FGM copula $C(u,v) = uv\{1+\alpha(1-u)(1-v)\}$, $-1 \leq \alpha \leq 1$, where the association parameter α is related to Kendall's τ rank correlation coefficient and the relation is expressed as $\tau = \frac{2}{9}\alpha$, which implies that $\tau \in [-0.22, 0.22]$. This copula was used in [13] for modeling the flood peak and flood volume and Kendall's τ was fitted as $\widehat{\tau} = 0.162$, so $\widehat{\alpha} = 0.73$.*

Considering the same model in [13], we show the influence of Kendall's τ on pmf values of $(T(3), S(2))$ given in Proposition 2.9 and the results are presented in the following table. From Table 3, it can be said that the pmf values of $(T(3), S(2))$ are similar for $\tau = 0$ and $\widehat{\tau} = 0.162$. As association parameter α increases, τ increases and $\tau = 0$ at $\alpha = 0$ corresponds the independence case.

Table 3. Comparison of pmf values of $(3,2)$nd upper record time vector for FGM copula with corresponding Kendall's τ rank correlation coefficient.

Prob.	$\tau_{\min} = -0.22$	$\tau = 0$	$\widehat{\tau} = 0.162$	$\tau_{\max} = 0.22$
$P\{T(3) < S(2)\}$	0.177	0.153	0.149	0.119
$P\{T(3) > S(2)\}$	0.821	0.798	0.799	0.820
$P\{T(3) = S(2)\}$	0.002	0.049	0.052	0.061

Below, we provide the graph of pmf values of the $(3,2)$th upper record time vector versus Kendall's τ rank correlation coefficient between flood peak and flood volume.

From Figure 3 it can be observed that in the case of $k < j$, the probability $P\{T(3) < S(2)\}$ decreases as τ changes from -0.22 to 0.22, while in the case of $k = j$ the probability $P\{T(3) = S(2)\}$ increases. If $k > j$, the probability $P\{T(3) > S(2)\}$ decreases as τ changes from -0.22 to 0 (negative dependence decreases) and this probability increases as τ changes from 0 to 0.22. Let us recall that if α changes from -1 to 0, then the negative dependence between X and Y increases, and if α changes from 0 to 1, the positive dependence increases. As a result, we can say that the FGM copula is suitable for analyzing the joint pmf of the $(3,2)$nd record time vector in flood data when there is a weak dependence between flood peak and volume as given in [13] with $\widehat{\tau} = 0.162$ since $\tau \in [-0.22, 0.22]$.

Considering the extended sequence $\{(X_k, Y_l)\}_{k\geq 1, l\geq 1}$, we are interested in the joint cdf of the marginal record values in the following section.

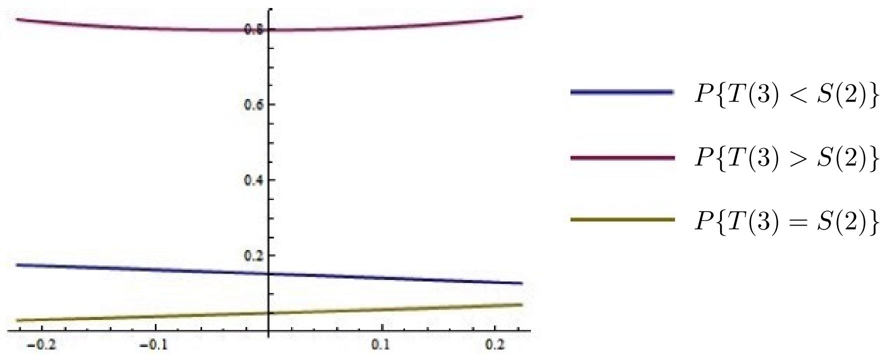

Figure 3. The graph of probabilities given in Proposition 3 of FGM copula plotted versus Kendall's τ rank correlation coefficient.

2.2. Joint cumulative distribution function of marginal records

It is well known that the marginal cdf of the second upper record value of the sequence $\{X_k\}_{k\geq 1}$ is

$$P\{X_{T(2)} \leq t\} = [1 - F_X(t)]\ln[1 - F_X(t)] + F_X(t) \tag{2.6}$$

and the corresponding pdf is

$$f_2(t) = -\ln[1 - F_X(t)]f(t). \tag{2.7}$$

The joint distribution function of the $(2,2)$nd upper record value vector from the extended bivariate sequence $\{(X_k, Y_l)\}_{k\geq 1, l\geq 1}$ is given in the following theorem.

Theorem 2.12 *Let us consider the sequence of iid bivariate random vectors $\{(X_k, Y_k)\}_{k\geq 1}$ with absolutely continuous cdf $F(x,y)$ and the joint pdf $f(x,y)$. Assume that $F_X(x)$ and $F_Y(y)$ are strictly increasing marginal cdfs. Consider the upper record value vector $(X_{T(2)}, Y_{S(2)})$ in the extended sequence $\{(X_k, Y_l)\}_{k\geq 1, l\geq 1}$. Then the joint cdf of the $(2,2)$nd upper record value vector is*

$$\begin{aligned}
&P\{X_{T(2)} \leq t, Y_{S(2)} \leq s\} \\
&= \int_{-\infty}^{s}\int_{-\infty}^{t} \Bigg\{ [F(t,y) - F(x,y)] \frac{[F_Y(s) - F_Y(y)]}{[1 - F_Y(y)]} \\
&+ [F(x,y) - F(t,y) - F(x,s) + F(t,s)] \\
&+ [F(x,s) - F(x,y)] \frac{[F_X(t) - F_X(x)]}{[1 - F_X(x)]} \Bigg\} \frac{1}{[1 - F(x,y)]} dF_{XY}(xy).
\end{aligned}$$

Proof The proof can be followed from the fact that

$$\begin{aligned}
&P\{X_{T(2)} \leq t, Y_{S(2)} \leq s\} \\
&= \sum_{i=2}^{\infty}\sum_{j=2}^{\infty} P\{X_{T(2)} \leq t, T(2) = i, Y_{S(2)} \leq s, S(2) = j\} \\
&= \sum_{i=2}^{\infty}\sum_{j=2}^{\infty} P\{X_i \leq t, X_1 > X_2, ..., X_1 > X_{i-1}, X_1 < X_i, Y_j \leq s, Y_1 > Y_2, ..., Y_1 > Y_{j-1}, Y_1 < Y_j\},
\end{aligned}$$

where the summation $\sum_{i=2}^{\infty}\sum_{j=2}^{\infty}$ is divided into three terms: $\sum_{i=2}^{\infty}\sum_{j=i+1}^{\infty}$ for $i<j$, $\sum_{j=2}^{\infty}\sum_{i=j+1}^{\infty}$ for $i>j$, and $\sum_{i=2}^{\infty}$ for $i=j$. By conditioning upon $X_1=x, Y_1=y$ and applying the total probability law, the proof is obtained. □

Corollary 2.13 *The joint pdf $f_{2,2}(t,s)$ of the vector $(X_{T(2)}, Y_{S(2)})$ can be derived by differentiating the cdf given in Theorem 2.12 with respect to t and s:*

$$f_{2,2}(t,s)$$
$$= \int_{-\infty}^{s}\int_{-\infty}^{t} \left\{ \frac{F^{1,0}(t,y) f_Y(s)}{[1-F_Y(y)]} + \frac{F^{0,1}(x,s) f_X(t)}{[1-F_X(x)]} + f(t,s) \right\} \frac{f(x,y)}{[1-F(x,y)]} dxdy,$$

where $F^{1,0}(t,y) = \frac{\partial F(t,y)}{\partial t}$, $F^{0,1}(x,s) = \frac{\partial F(x,y)}{\partial s}$ are the partial derivatives of the random vector (X,Y) and $f_X(t) = \frac{dF_X(t)}{dt}$, $f_Y(s) = \frac{dF_Y(s)}{ds}$ are the marginal pdfs of X and Y random variables, respectively.

Example 2.14 *Let us consider that X and Y have bivariate FGM distribution where marginal distributions are both standard exponential, i.e. $F(x,y) = (1-e^{-x})(1-e^{-y})(1+\alpha e^{-x-y})$ $x,y>0$ and $-1 \le \alpha \le 1$. FGM distribution has a simplificative analytical form and is convenient for calculations. Since the expression of the cdf of $(X_{T(2)}, Y_{S(2)})$ is unwieldy, in the following table we provide some numerical results of the cdf of $(X_{T(2)}, Y_{S(2)})$ in the case of the underlying distribution being FGM when the association parameter is $\alpha \neq 0$, also including the independence case, i.e. $\alpha = 0$.*

According to the Table 4, it can be seen that if association parameter α increases from 0 to 1, then the probability $P\{X_{T(2)} \le t, Y_{S(2)} \le s\}$ decreases.

In the next theorem, we provide the joint cdf of $X_{T(2)}, Y_{S(2)}$, and $X_{T(3)}$.

Theorem 2.15 *The joint cdf of the upper record value vector $(X_{T(2)}, Y_{S(2)}, X_{T(3)})$ is*

$$P\{X_{T(2)} \le t, Y_{S(2)} \le s, X_{T(3)} \le r\}$$
$$= \int_{-\infty}^{t}\int_{-\infty}^{s}\int_{-\infty}^{z} \left\{ \frac{[F_Y(y)[F(z,s)-F(z,y)][F_X(r)-F_X(z)]}{[1-F(z,y)][1-F_X(z)]} + \frac{[F_Y(y)[F(r,y)-F(z,y)][F_Y(s)-F_Y(y)]}{[1-F(z,y)][1-F_Y(y)]} \right.$$
$$+ \frac{[F_Y(y)[F(z,y)-F(r,y)-F(z,s)+F(r,s)]}{[1-F(z,y)]} + \frac{[F(x,s)-F(x,y)][F_X(r)-F_X(z)]}{[1-F_X(x)][1-F_X(z)]}$$
$$\left. + \frac{[F_Y(s)-F_Y(y)]\ [F_X(r)-F_X(z)]}{[1-F_X(z)]} \right\}$$
$$\times \frac{1}{[1-F(x,y)]} f_X(z) f(x,y) dxdydz.$$

Proof The proof is similar to that of Theorem 2.12. □

3. Conclusion

In this paper, we introduce the marginal record values and record times in extended sequences of bivariate random vectors. Defining a record in bivariate or multivariate random sequences is more complicated since

Table 4. The cdf values of $P\{X_{T(2)} \leq t, Y_{S(2)} \leq s\}$ for FGM distribution with standard exponential marginals with different values of α.

	$\alpha = 0$						
t / s	2	3	4	5	10	15	25
2	0.3528	0.4757	0.5396	0.5699	0.5937	0.5940	0.5940
3	0.4757	0.6414	0.7275	0.7685	0.8000	0.8001	0.8001
4	0.5396	0.7275	0.8252	0.8717	0.9080	0.9084	0.9084
5	0.5699	0.7685	0.8717	0.9208	0.9591	0.9596	0.9596
10	0.5937	0.8000	0.9080	0.9591	0.9990	0.9995	0.9995
15	0.5940	0.8001	0.9084	0.9596	0.9995	0.9999	0.9999
25	0.5940	0.8001	0.9084	0.9596	0.9995	0.9999	0.9999

	$\alpha = 0.5$						
t / s	2	3	4	5	10	15	25
2	0.3065	0.2900	0.2602	0.2296	0.4791	0.4792	0.4792
3	0.2900	0.2735	0.2438	0.2132	0.6143	0.6144	0.6144
4	0.2602	0.2438	0.2141	0.1835	0.6769	0.6771	0.6771
5	0.2296	0.2132	0.1835	0.1528	0.7039	0.7041	0.7041
10	0.4791	0.6143	0.6769	0.7039	0.7223	0.7225	0.7225
15	0.4792	0.6144	0.6771	0.7041	0.7225	0.7227	0.7227
25	0.4792	0.6144	0.6771	0.7041	0.7225	0.7227	0.7227

	$\alpha = 1$						
t / s	2	3	4	5	10	15	25
2	0.2764	0.3531	0.3877	0.4023	0.4121	0.4122	0.4122
3	0.3531	0.4461	0.4878	0.5053	0.5170	0.5171	0.5171
4	0.3877	0.4878	0.5325	0.5512	0.5636	0.5637	0.5637
5	0.4023	0.5053	0.5512	0.5704	0.5832	0.5833	0.5833
10	0.4121	0.5170	0.5636	0.5832	0.5961	0.5962	0.5962
15	0.4122	0.5171	0.5637	0.5833	0.5962	0.5963	0.5963
25	0.4122	0.5171	0.5637	0.5833	0.5962	0.5963	0.5963

vectors in n-dimensional space are partially ordered. Therefore, only a few papers in the literature deal with bivariate and multivariate records. In the present paper, the joint pmfs of some upper record time vectors are derived and these pmfs are shown to be marginal free. Furthermore, the joint cdfs and pdfs of some upper record value vectors are provided. The obtained pmfs of record times and cdfs of record values come into prominence while predicting future records based on past observations. Assuming records in rainfall intensity and rainfall depth as an example of bivariate record data, we can predict the next record value of rainfall depth given the record value of rainfall intensity having observations up to the present time. This prediction is crucial for reducing the risk and preventing extreme flooding events. Furthermore, the records of bivariate sequences naturally arise in many fields of applications of probability and statistics. Example 2.11 presents a case where bivariate records are important in flood frequency analysis, where the variables of interest are flood volume and flood peak. The study of distributional properties of bivariate records of these variables allows us to calculate predicted values for future flooding, which is extremely important in the design and construction of dams to prevent damage that can be caused by flooding. We believe that the results obtained in this paper will be motivating for developing the theory of bivariate and multivariate records. In statistical inference, the results may be useful for estimating association parameters of underlying distributions when we need to use bivariate record data, since the MLE estimators require the joint pmf of bivariate records. Below, we discuss some open problems that may be important for future research in this area.

Finding the joint distributions of marginal records for general n and m is a challenging problem requiring complex technical calculations. The derivation of the joint distribution of marginal record values and record times for any n and m is still an open problem. It is also interesting to investigate under which conditions the (n,m)th record value vector and record time vector possess the Markov property.

References

[1] Ahsanullah M. Characterization of the exponential distribution by record values. Sankhya Series B 1979; 116-121.

[2] Ahsanullah M. Record values of independent and identically distributed continuous random variables. Pakistan Journal of Statistics 1992; 8: 9-34.

[3] Ahsanullah M. Some inferences of uniform distribution based on record values. Pakistan Journal of Statistics 1994; 7: 27-33.

[4] Ahsanullah M. Record Statistics. New York, NY, USA: Nova Science Publishers, 1995.

[5] Ahsanullah M, Holland B. Record values and the geometric distribution. Statistical Papers 1984; 25: 319-327.

[6] Arnold BC, Balakrishnan N, Nagaraja HN. Records. Hoboken, NJ, USA: Wiley, 1998.

[7] Bayramoglu I. On the records of multivariate random sequences. Metrika 2016; 79: 725-747.

[8] Benestad RE. How often can we expect a record event?. Climate Research 2003; 25: 3-13.

[9] Benestad RE. Record-values, nonstationarity tests and extreme value distributions. Global Planet Change 2004; 44: 11-26.

[10] Blokhinov YG, Sarmanov O. Gamma correlation and its use in computations of long-term streamflow regulation. Soviet Hydrology Selected Papers 1968; 1: 36-54.

[11] Chandler KN. The distribution and frequency of record values. Journal of the Royal Statistical Society 1952; 220-228.

[12] Clarke RT. Bivariate gamma distributions for extending annual streamflow records from precipitation: some large-sample results. Water Resources Research 1980; 16: 863-870.

[13] Favre AC, El Adlouni S, Perreault L, Thiémonge N, Bobée B. Multivariate hydrological frequency analysis using copulas. Water Resources Research 2004; 40: W01101.

[14] Gnedin AV. Conical extremes of a multivariate sample. Journal of Research of the National Institute of Standards and Technology 1994; 99: 511-511.

[15] Gnedin AV. Records from a multivariate normal sample. Statistics & Probability Letters 1998; 39: 11-15.

[16] Goldie CM, Resnick S. Records in a partially ordered set. Annals of Probability 1989; 678-699.

[17] Goldie CM, Resnick SI. Many multivariate records. Stochastic Processes and Their Applications 1995; 59: 185-216.

[18] Gumbel EJ, Mustafi CK. Some analytical properties of bivariate extremal distributions. Journal of the American Statistical Association 1967; 62: 569-588.

[19] Nagaraja HN. Record values and related statistics-a review. Communications in Statistics - Theory and Methods 1988; 17: 2223-2238.

[20] Nevzorov VB. Records. Theory of Probability and its Applications 1988; 32: 201-228.

[21] Nevzorov VB. Records: Mathematical Theory. Translation of Mathematical Monographs, Vol. 194. Providence, RI, USA: American Mathematical Society, 2001.