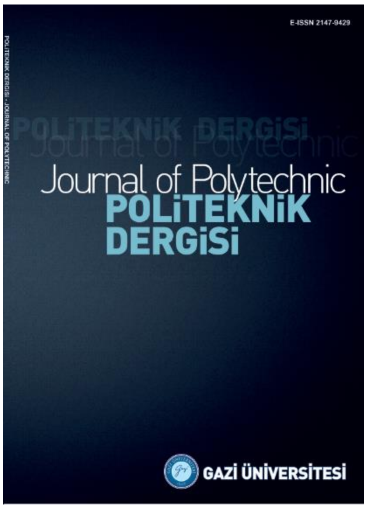

POLİTEKNİK DERGİSİ

JOURNAL of POLYTECHNIC

ISSN: 1302-0900 (PRINT), ISSN: 2147-9429 (ONLINE)

URL: http://dergipark.org.tr/politeknik

A comparison of copula families on dependence structure of extreme order statistics

Uç sıra istatistiklerinin bağımlılık yapısı üzerine kapula ailelerinin bir karşılaştırılması

***Yazar(lar) (Author(s))**: Ferhan BAŞ KAMAN[1], Hülya OLMUŞ[2]*

ORCID[1]: 0000-0002-1879-9215

ORCID[2]: 0000-0002-8983-708X

Bu makaleye şu şekilde atıfta bulunabilirsiniz(To cite to this article): Baş Kaman F. ve Olmuş H., "Uç sıra istatistiklerinin bağımlılık yapısı üzerine kapula ailelerinin bir karşılaştırılması", ***Politeknik Dergisi***, 22(4): 855-862 (2019).

Erişim linki (To link to this article): http://dergipark.org.tr/politeknik/archive

DOI: **10.2339/politeknik.421978**

A Comparison of Copula Families on Dependence Structure of Extreme Order Statistics

Araştırma Makalesi / Research Article

Ferhan BAŞ KAMAN[1*], Hülya OLMUŞ[2]
[1]Şereflikoçhisar Faculty of Applied Science, Department of Banking and Finance, Yıldırım Beyazıt University, Ankara, Turkey
[2] Faculty of Science, Department of Statistics, Gazi University, Ankara, Turkey

(Geliş/Received : 08.05.2018 ; Kabul/Accepted : 02.10.2018)

ABSTRACT

Order statistics have an important place in probability theory and statistical inference, especially reliability analysis. It is aimed to compare the dependency structure with some copula families of min-max copula for the extreme order statistics in this study. Suitability of Clayton, Frank, Gumbel from Archimedean copula families, Gaussian, Farlie-Gumbel-Morgenstern and min-max copula are examined by simulation study. To find the most suitable copula family, some model selection criteria are used and important inferences are obtained.

Keywords: Dependent structure, order statistics, copula, simulation.

Uç Sıra İstatistiklerinin Bağımlılık Yapısı Üzerine Kapula Ailelerinin Bir Karşılaştırılması

ÖZ

Sıra istatistikleri, olasılık teorisinde ve istatistiksel çıkarımlarda özellikle güvenilirlik analizinde önemli bir yere sahiptir. Bu çalışmada uç sıra istatistiklerinin bağımlılık yapısı için min-max kapula ile bazı kapula aileleri karşılaştırılmıştır. Arşimedyen kapulalardan Clayton, Frank, Gumbel kapula aileleri, Gaussian, Farlie-Gumbel-Morgenstern ve min-max kapulaların uygunluğu simülasyon çalışması ile incelenmiştir. En uygun kapula ailesini bulmak için bazı model seçim kriterleri kullanılarak önemli çıkarımlar elde edilmiştir.

Anahtar Kelimeler: Bağımlılık yapısı, sıra istatistikleri, kapula, simülasyon.

1. INTRODUCTION

Order statistics commonly used in statistical hypotheses, estimation problems, statistical process controls, reliability and risk management and in many practical fields and therefore it is of great importance by researchers [1]. Developments in detailed theories and applications with respect to order statistics can be referred by Arnold et al., Balakrishnan and Rao and David and Nagaraja [2-5].

Let $X_{(1)} \le X_{(2)} \le ... \le X_{(n)}$ denote the order statistics of independent and identically distributed (i.i.d.) random variables $X_{1,} X_2, ..., X_n,\ n \in N$ with common distribution function F. It is known that, the lifetimes of series and parallel and systems are correspond to the $X_{(1)}$ and $X_{(n)}$, respectively. Therefore, many researchers have studied the dependency structure of extreme order statistics $X_{(1)}$ and $X_{(n)}$. One of the most important of these researches is min-max copula proposed by [6]. He found the dependency structure between $X_{(1)}$ and $X_{(n)}$ extreme order statistics where $X_{1,} X_2, ..., X_n,\ n \in N$, are i.i.d. random variables.

*Sorumlu Yazar (Corresponding Author)
e-posta : fbaskaman@ybu.edu.tr

Hence, Kendall's τ and Spearman's ρ are easily calculated and it is shown that $3\tau_n \ge \rho_n \ge \tau_n > 0$ and also $\frac{\rho_n}{\tau_n} \to \frac{3}{2}$ is found and numerically proven. It is proved Schmitz's assumption by applying L'Hospital's Rule to the copula between $X_{(1)}$ and $X_{(n)}$ by Li and Li and lastly studied three new formulas for Spearman's ρ between two identically distributed extreme order statistics by Chen, $X_{(1)}$ and $X_{(n)}$ when $n \ge 2$ by taking advantage of these results [7, 8]. Ghalibaf examined some measures of association between extreme order statistics $X_{(1)}$ and $X_{(n)}$. He showed that dependence is decreasing function form n [9]. The dependency structure of order statistics are also frequently studied using Archimedean copulas. Li and Fang examined ordering properties on order statistics from heterogeneous observations by using Archimedean copulas [10]. Mesfioui et al. are devoted to characterize several ordering properties of the maximum order statistics of heterogenous random variables with an Archimedean copula [11].

In order to investigate the relationship between variables, it is necessary to reveal the dependency structure between variables and because of this reason, the concept of "copula" is being used in statistics literature in recent years. Many researchers studied the dependency structure of order statistics by means of copula function. The aim of this study is to compare the dependency structure with some copula families of min-max copula for the extreme order statistics. In addition, using the rank based methods, important inferences has been made for copula families. To achieve this goal, this article is organised as follows: In section 2, basic properties of copulas are given. Also, copulas, including Clayton, Frank, Gumbel, Gaussian, Farlie-Gumbel-Morgenstern and min-max copula, are briefly introduced. In section 3, dependency and ranks are discussed and also it is tried to explain why rank-based methods are used. In section 4, model selection criteries are presented. In section 5, simulations of copula are presented and lastly, the results are discussed in section 6.

2. BASIC PROPERTIES OF COPULAS

A copula is a function $C:[0,1]^2 \to [0,1]$ which satisfies following properties [12].

1. For every $u,v \in I$;

$C(u,0) = C(0,v) = 0.$

$C(u,1) = u$ ve $C(1,v) = v.$

2.For every $u_1,u_2,v_1,v_2 \in I$ such that $u_1 \le u_2$ and $v_1 \le v_2$;

$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \ge 0,$

Following result is an important theory for copulas.

Theorem 1. (Sklar's Theorem) H is a joint distribution function with margins F and G. For every $x,y \in \overline{R}$;

$$H(x, y) = C(F(x), G(y)), \quad (1)$$

where C is a copula. If F and G are continuous, C is unique; otherwise, C is uniquely determined on $RangeF \times RangeG$. Contrarily, if C is a copula and F and G are distribution functions, then the function H defined by (1) is a joint distribution function with margins F and G. If F and G are continuous C is defined as follows [12].

$$C(u,v) = H(F^{(-1)}(u), G^{(-1)}(v)) \quad (2)$$

2.1. Kendall's Tau (τ)

Kendall's τ is a nonparametric measures of association of between two random variables X and Y. The relation of Kendall's τ with C copula is given by Theorem 2.

Theorem 2. Let C is common copula of continuous random variables X and Y. Then Kendall's τ is given as follows [12].

$$\tau(C) = 4\int_0^1\int_0^1 C(x,y)dC(x,y) - 1 \quad (3)$$

2.2. Archimedean Copulas

Archimedean copulas are widely using in applications because of they can be constructed easily and they have many nice algebraic properties. Archimedean copulas at first appeared not in statistics, but rather in the study of probabilistic metric spaces, where they were studied as part of the development of a probabilistic version of the triangle inequality [12].

Properties of Archimedean copulas

Let φ is a continous, strictly decreasing function on $\varphi: I \to [0,\infty]$ such that $\varphi(1)=0$ and $\varphi^{[-1]}$ is pseudo-inverse of φ and $Domain\,\varphi^{[-1]} = [0,\infty]$, $Range\,\varphi^{[-1]} = [0,1]$ is given by

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t) & ,0 \le t \le \varphi(0) \\ 0 & ,\varphi(0) \le t \le \infty \end{cases} \quad (4)$$

Point out that $\varphi^{[-1]}$ is a continous and nonincreasing on $[0,\infty]$ and strictly decreasing on $[0,\varphi(0)]$. Hence, $C:[0,1]^2 \to [0,1]$ is defined as follows:

$$C(u,v) = \varphi^{[-1]}(\varphi(u)+\varphi(v)) \quad (5)$$

This form (5) of copulas are called Archimedean copulas and φ is called a generator of the copula. If $\varphi(0)=\infty$, φ is a strict generator and in this case, $\varphi^{[-1]}(\varphi(u)+\varphi(v))$ is said to be a strict Archimedean copula. Properties of generators are given as below.

1. $\varphi(1) = 0$.
2. For every $t \in (0,1)$, $\varphi'(t) < 0$.
3. For every $t \in (0,1)$, $\varphi''(t) \ge 0$.

Theorem 3. Let C is an Archimedean copula and φ its generator;

1. C is symmetric, namely, for every $u,v \in I$, $C(u,v) = C(v,u)$.
2. C is associative, namely, for every $u,v,w \in I$, $C(C(u,v),w) = C(u,C(v,w))$.

Table 1. Archimedean copulas used in this study

Copula	$C_\theta(u,v)$	$\varphi_\theta(t)$	Parameter interval	Kendall's τ
Clayton	$[max(u^{-\theta}+v^{-\theta}-1,0)]^{-\frac{1}{\theta}}$	$\frac{1}{\theta}(t^{-\theta}-1)$	$[-1,\infty)\setminus\{0\}$	$\frac{\theta}{\theta+2}$
Frank	$-\frac{1}{\theta}\ln(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{(e^{-\theta}-1)})$	$-\ln\frac{e^{-\theta t}-1}{e^{-\theta}-1}$	$(-\infty,\infty)\setminus\{0\}$	$1-\frac{4}{\theta}[D_1(-\theta)-1]$
Gumbel	$\exp(-[(-\ln u)^\theta+(-\ln v)^\theta]^{\frac{1}{\theta}})$	$(-\ln t)^\theta$	$[1,\infty)$	$\frac{\theta-1}{\theta}$

*where $D_1(\theta) = \frac{1}{\theta}\int_0^\theta \frac{t}{e^t-1}$

3. If $c > 0$ is any constant, $c\varphi$ is also a generator of C [12].

Table 1 gives the Archimedean copulas which used in this study, their functions $C_\theta(u,v)$, generators $\varphi_\theta(t)$, parameter intervals and connecting with Kendall's τ [12]

2.3. Farlie-Gumbel-Morgenstern (FGM) Copula

FGM copula is defined as follows:

$$C_\theta(u,v) = uv[1+\theta(1-u)(1-v)], \quad \theta \in [-1,1] \qquad (6)$$

The relations of Kendall's τ with FGM is $\tau = \frac{2\theta}{9}$ [12].

2.4. Gaussian Copula

The Gaussian copula is derived from multivariate Gaussian distribution. Φ is the distribution function of one-dimensional standard normal distribution and let Φ_Σ^n is the distribution function of the positive definite correlation matrix Σ and n-dimensional standard normal distribution. Hence, n-dimensional Gaussian copula C_Σ^Φ is defined as follows;

$$C_\Sigma^\Phi(u_1,...,u_n) = \Phi_\Sigma^n(\Phi^{-1}(u_1),...,\Phi^{-1}(u_n)), \quad \forall(u_1,...,u_n) \in [0,1]^n \qquad (7)$$

For $n = 2$, 2-dimensional Gaussian copula is defined as follows:

$$C_{\rho_{12}}^\Phi(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho_{12}{}^2)^{1/2}} \exp\left(-\frac{s^2 - 2\rho_{12}\cdot s\cdot t + t^2}{2(1-\rho_{12}{}^2)}\right) ds\,dt, \quad (u,v) \in [0,1]^2 \qquad (8)$$

where ρ_{12} is the correlation coefficient of bivariate standard normal distribution. The relation of Kendall's τ with Gaussian copula is defined as $\tau = \frac{2}{\pi}\arcsin(\rho_{12})$ [13].

2.5. Min-max Copula

$X_{(1)}$ and $X_{(n)}$ denote the extreme order statistics of i.i.d. random variables $X_1, X_2, ..., X_n$, $n \in N$. The marginals of $X_{(1)}$ and $X_{(n)}$ and the joint cumulative distribution function of $X_{(1)}$ and $X_{(n)}$ are given as follows [2].

$$F_1(x) = P\{X_{(1)} \le x\} = 1-[1-F(x)]^n, \quad -\infty < x < \infty \qquad (9)$$

$$F_n(x) = P\{X_{(n)} \le x\} = F(x)^n, \quad -\infty < x < \infty \qquad (10)$$

$$F_{(1),(n)}(x,y) = P\{X_{(1)} \le x, X_{(n)} \le y\} = \begin{cases} F(y)^n - (F(y)-F(x))^n, & x < y \\ F(y)^n, & x \ge y \end{cases} \qquad (11)$$

The copula can be derived from $C(u,v) = H(F^{(-1)}(u), G^{(-1)}(v))$ where $F^{(-1)}$ denotes the generalized inverse. Here $F^{(-1)}(u)$ and $G^{(-1)}(v)$ are found as follows:

$$u = 1-[1-F(x)]^n, \quad F^{(-1)}(u) = 1-[1-u]^{1/n}, \qquad v = G(y)^n, \quad G^{(-1)}(v) = v^{1/n}$$

Then, by solving the equation $C_n(F_1(x), F_n(y)) = F_{(1),(n)}(x,y)$

$$C_n = \begin{cases} v - (v^{1/n} + (1-u)^{1/n} - 1)^n, & 1-(1-u)^{1/n} < v^{1/n} \\ v, & 1-(1-u)^{1/n} \ge v^{1/n} \end{cases} \qquad (12)$$

C_n is call min-max-copula and it describes the dependence between $X_{(1)}$ and $X_{(n)}$ completely [6].

3. DEPENDENCE AND RANKS

Let assume that $X_{11}, X_{12}, ...X_{1n}$ and $X_{21}, X_{22}, ...X_{2n}$ are given from pairs (X_1, X_2) of continuous variables. $H(x_1, x_2)$ is bivariate distribution function of (X_1, X_2). In view of Sklar's theorem, there exists a unique copula C for which identity, Eq. (1), holds. Thus, by using $F(x_1)$ and $G(x_2)$, uniquely C copula can be derived. Consider the transformed pairs

$Z_i = \exp(2X_{1i})$, $T_i = \exp(3X_{2i})$.

The pairs (Z,T) are transformed of the pairs (X_1, X_2) and both of them have same copula [14].

Table 2 shows five independent pairs of mutually independent observations (X_{1i}, X_{2i}) generated from the $N(2,1)$ distribution using Matlab R2017b

Table 2. Sample data

i	1	2	3	4	5
X_{1i}	0.4612	2.9783	0.4443	2.2599	3.2763
X_{2i}	1.6727	3.8142	1.8476	1.3407	2.1501

In Figure 1, scatter plots of the pairs (X_1, X_2), transformed data (Z,T) and ranks derived from the (X_1, X_2) and (Z,T) are given.

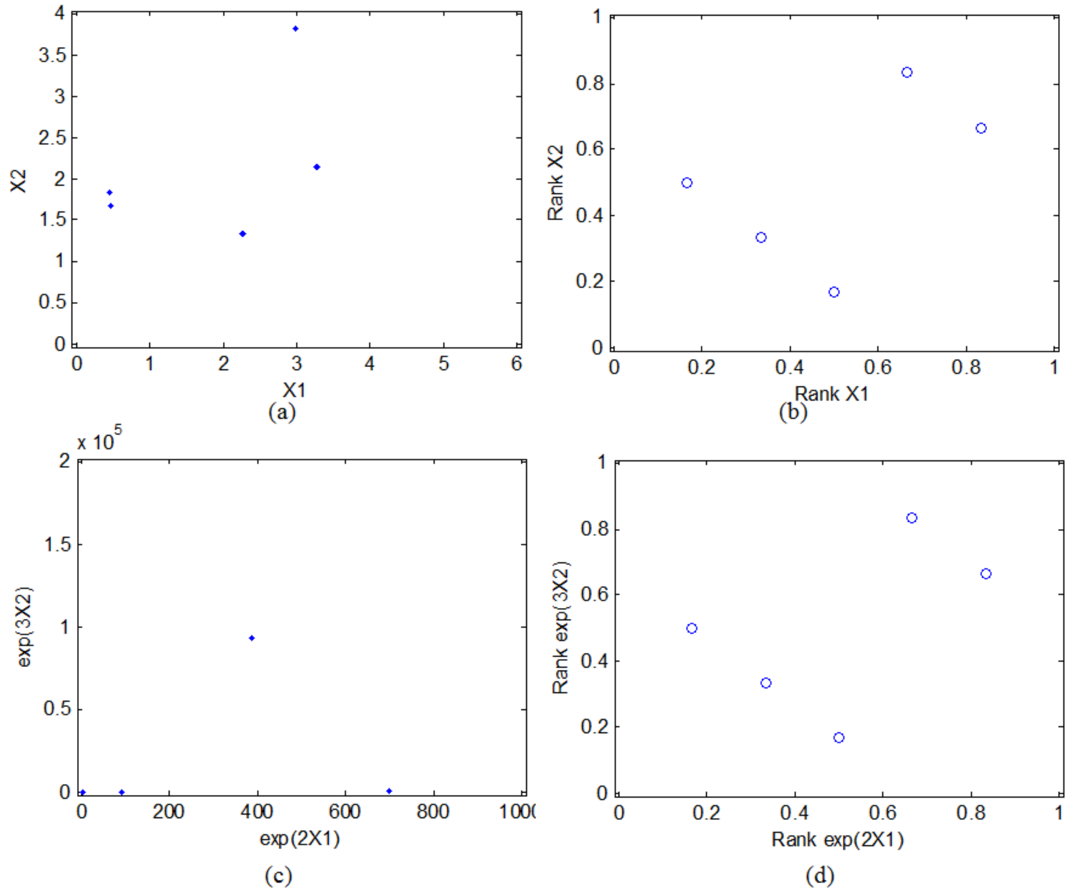

Figure 1. Scatter plots of the pairs (a) (X_1, X_2) (b) $(Z,T)=(\exp(2X_{1i}),\exp(3X_{2i}))$ (c) ranks of (X_1, X_2) and (d) ranks of $(Z,T)=(\exp(2X_{1i}),\exp(3X_{2i}))$

When Figure 1 is examined, scatter plots of (X_1, X_2) and transformed data (Z,T) are substantially different, however, scatter plots of ranks of (X_1, X_2) and transformed data (Z,T) are completely same. Therefore, the use of ranks lead to more accurate results. In this study, rank based methods are used for copulas.

Let (X,Y) be measured data, u_i, v_i are rank of (X,Y) which are defined as

$$\begin{cases} u_i = \dfrac{rank(X_i)}{N+1} \\ v_i = \dfrac{rank(Y_i)}{N+1} \end{cases} \quad i=1,2,...,N \tag{13}$$

where $rank(X_i)$(or $rank(Y_i)$) denotes the rank of X_i (or Y_i) among X (or Y) in an ascending order. Thus the measured data in original space is transformed into the standard uniform random vector (u,v) [15].

4. MODEL SELECTION CRITERIES

There are many methods in the literature to choose the most suitable copula for the database. First, Chi-square goodness of fit test is determined as suitability of copulas for dataset. Then the best-fit copula is obtained by Akaike Information Criterion and Bayesian Information Criterion and lastly, scatter plot of generating data from copulas are examined.

Chi-square goodness of fit: Chi-square goodness of fit test is calculated as follows:

$$\chi^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(O_{ij}-E_{ij})^2}{E_{ij}} \tag{14}$$

Here O_{ij} represents the observed frequencies and E_{ij} represents the expected frequencies. Degrees of freedom is calculated as $df=((I-1)(J-1))-p-(q-1)$, I is the number of rows, J is the number of columns, p is the number of parameters estimated and q is the number of cells pooled together. If the table value corresponding to the degrees of freedom is greater than the calculated chi-square, the copula is found suitable [16].

Akaike Information Criterion- AIC and Bayesian Information Criterion- BIC : The best-fit copula is identified by comparing the evaluated values of AIC and BIC which are defined as

$$AIC = -2\sum_{i=1}^{N} \ln c(u_i, v_i) + 2k, \qquad (15)$$

$$BIC = -2\sum_{i=1}^{N} \ln c(u_i, v_i) + k \ln N \qquad (16)$$

where k is the number of copula parameters, N is the sample size and $c(u_i, v_i)$ is the probability density function of $C(u_i, v_i)$, (u_i, v_i) obtained from Eq. (13) [17, 18]. A copula corresponding to the lowest AIC value or BIC value is determined as the best-fit copula [15].

Scatter Plot: The scatter plot of ranks of original data can be also capture the dependency structure. The most similar scatter plot of data obtained from copula to the scatter plot of the ranks of original data is determined as best-fit copula [14].

5. SIMULATION STUDY

In this section, suitability of Clayton, Frank, Gumbel from Archimedean copula families, Gaussian, FGM and min-max copula are examined by simulation study. Avérous et al. formalized the proof that the copula associated with a pair of order statistics does not depend on the parent distribution [19]. Based on this study, firstly, we drawn random samples from normal, uniform and exponential distributions. After we have seen that similar results were obtained from all distributions we use, we only used normal distribution. Let's assume random samples are drawn from $X \sim N(2,1)$ distribution and $X_{(1)}, X_{(2)}, ..., X_{(n)}$ are order statistics of these samples. The dependency relationships of $X_{(1)}, X_{(n)}$ order statistics for different n values $(n = 2, 5, 10, 20, 100, 200)$ are evaluated in this study. First, to obtain observed frequencies in Chi-square goodness of fit test, we used the ranks of $X_{(1)}, X_{(n)}$ order statistics classified with 4×4 quartile points (for $j = 1,2,3,4$ $1000 \times \frac{j}{4}$). In Table 3, τ value is Kendall's tau obtained from $X_{(1)}, X_{(n)}$ order statistics, θ are parameter estimates for copula families by using τ value. Then, expected frequencies are obtained from copulas by means of θ. It is repeated r=500 times. χ^2 values for each copula family is calculated by using the

Table 3. Comparison of simulation results for copulas

n	τ	Copula	θ	χ^2	$\chi^2_{8;.05}$	p-value	AIC	BIC
2	0.3294	Frank	3.2593	140.0462	15.51	0.0000	-230.4088	-225.5011
		Clayton	0.9826	116.5814	15.51	0.0000	-169.7322	-164.8254
		Gumbel	1.4913	118.3252	15.51	0.0000	-213.7415	-208.8337
		FGM	1.4825	-	15.51	-	-	-
		Min-max	2	14.3332	15.51	0.0735	-3377.5	-3372.6
		Gaussian	0.4947	114.8323	15.51	0.0000	-249.6951	-244.7874
5	0.0965	Frank	0.8758	22.0341	15.51	0.0049	-18.6516	-13.7438
		Clayton	0.2138	19.1601	15.51	0.0140	-11.5356	-6.6278
		Gumbel	1.1069	19.7345	15.51	0.0114	-10.8259	-5.9181
		FGM	0.4346	17.8609	15.51	0.0223	-19.7405	-14.8327
		Min-max	5	11.7187	15.51	0.1642	-929.211	-924.3033
		Gaussian	0.1511	17.4738	15.51	0.0255	-19.0115	-14.1037
10	0.0392	Frank	0.3532	12.0316	15.51	0.1498	-1.3707	3.5371
		Clayton	0.0816	11.4153	15.51	0.1793	0.2492	5.1567
		Gumbel	1.0408	10.8653	15.51	0.2094	-2.0630	2.8447
		FGM	0.1764	10.7895	15.51	0.2139	-1.3509	3.5569
		Min-max	10	12.9961	15.51	0.1120	-422.0772	-417.1695
		Gaussian	0.0615	11.2567	15.51	0.1876	-2.9461	1.9617
20	0.0259	Frank	0.2333	14.6592	15.51	0.0661	0.4755	5.3832
		Clayton	0.0532	14.3174	15.51	0.0739	-1.7111	3.1966
		Gumbel	1.0266	15.3692	15.51	0.0524	2.3511	7.2591
		FGM	0.1166	15.1108	15.51	0.0570	0.4891	5.3968
		Min-max	20	15.4470	15.51	0.0510	-203.4450	-198.5372
		Gaussian	0.0407	15.2036	15.51	0.0553	0.3343	5.2421
100	0.0135	Frank	0.1215	11.5544	15.51	0.1722	1.5878	6.4956
		Clayton	0.0274	12.1262	15.51	0.1457	0.6212	5.5289
		Gumbel	1.0137	12.0672	15.51	0.1482	1.8853	6.7930
		FGM	0.0607	11.5576	15.51	0.1721	1.5831	6.4908
		Min-max	100	11.9264	15.51	0.1545	-39.040	-34.1862
		Gaussian	0.0212	12.0399	15.51	0.1494	0.8045	5.7132
200	0.0062	Frank	0.0555	17.2599	15.51	0.0275	1.9103	6.8181
		Clayton	0.0124	17.3160	15.51	0.0270	1.8614	6.7691
		Gumbel	1.0062	17.7051	15.51	0.0235	2.0170	6.9247
		FGM	0.0277	17.0802	15.51	0.0293	1.9091	6.8169
		Min-max	200	17.4716	15.51	0.0256	-18.3476	-13.4398
		Gaussian	0.0097	17.5871	15.51	0.0245	1.9854	6.8932

obtained values. The below hypothesis is used to test the suitability of each and every family to data set.

H_0 : Copula family is suitable for dataset.

H_1: Copula family is not suitable for dataset.

Degrees of freedom is calculated via frequency matrix classified with 4×4 quartile points and the value of $\chi^2_{8;.05}=15.51$ is compared with calculated χ^2 values for all copula families. Then, the best-fit copula is determined by AIC and BIC. The ranks of frequency matrix of $X_{(1)}$, $X_{(n)}$ order statistics are again used in here. All situations and results are given in Table 3. Since the parameter range of the FGM copula and Kendall's τ value for FGM copula are between [-1,1] and [-2 / 9, 2 / 9], respectively, this copula is not applicable for $n=2$

When $n=2$ and $n=5$, we can not reject H_0 for min-max copula but we can reject H_0 for Clayton, Frank, Gumbel, Gaussian and FGM copula families. Therefore, only min-max copula is founded suitable for data set. For $n=2$, when we examine the highest AIC and BIC values, first eliminated copula is Clayton and later, Gumbel, Frank and Gaussian, respectively. For $n=5$, first eliminated copula is Gumbel and later, Clayton, Frank, Gaussian and FGM are determined.

When $n=10$, we can not reject H_0 for all copula families and all of them are founded suitable for data set. Here, the copula with the lowest AIC and BIC values is identified as min-max copula, hence, as the best-fit copula. Later, Gaussian, Gumbel, Frank, FGM and lastly the worst copula as Clayton are determined.

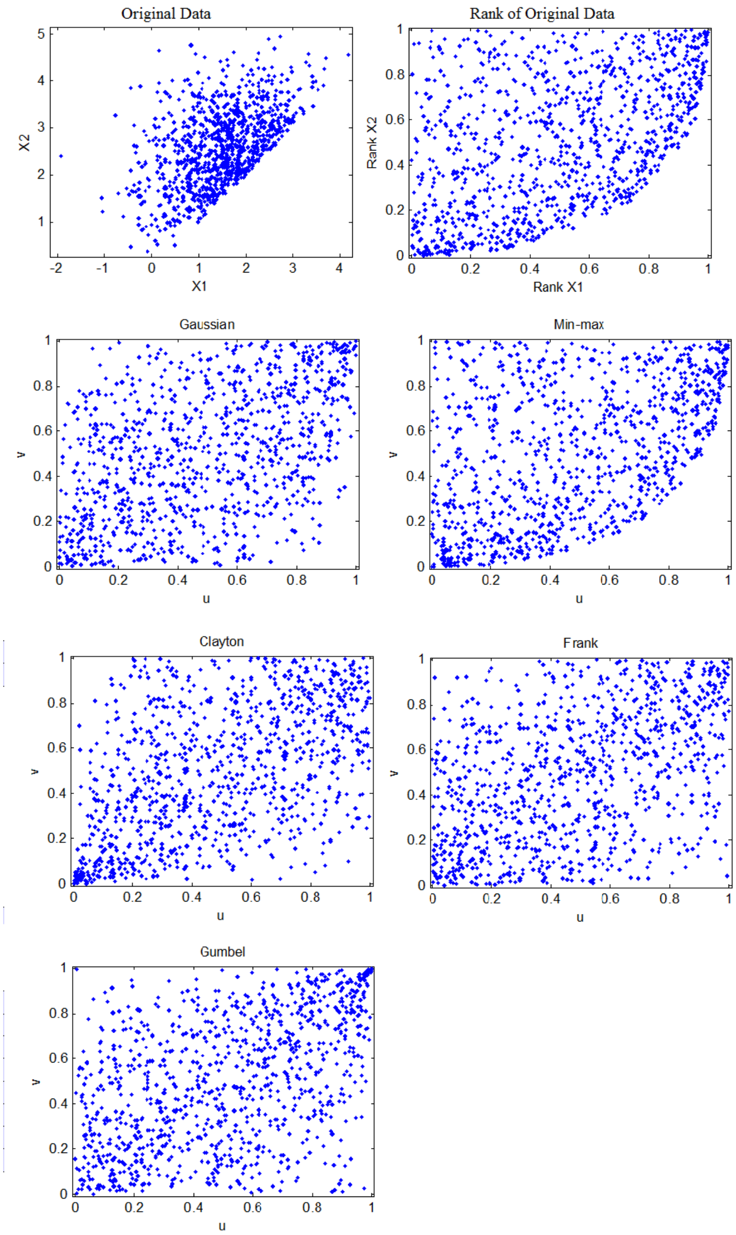

Figure 1. Scatter plots of original data, ranks of original data and copulas for $n=2$

When $n = 20$ and $n = 100$, we can not reject H_0 for all copula families and all of them are founded suitable for data set again. Here, the copula with the lowest AIC and BIC values is identified as min-max copula, hence, as the best-fit copula. Then, Clayton, Gaussian, FGM, Frank and Gumbel are identified, respectively.

When $n = 200$, we can reject H_0 for all copula families and all of them are not founded suitable for data set. Here, first eliminated copula is Gumbel and later, Gaussian, Frank, FGM, Clayton and Min-max copula. We can say that Min-max copula is better than others but AIC and BIC values of other copula families are very close to each other, so it wouldn't be right to say clearly which one is better.

The scatter plot of ranks of original data can be also capture the dependency structure.

Figure 1 and Figure 2 shows scatter plots for $n = 2$ and $n = 10$, respectively.

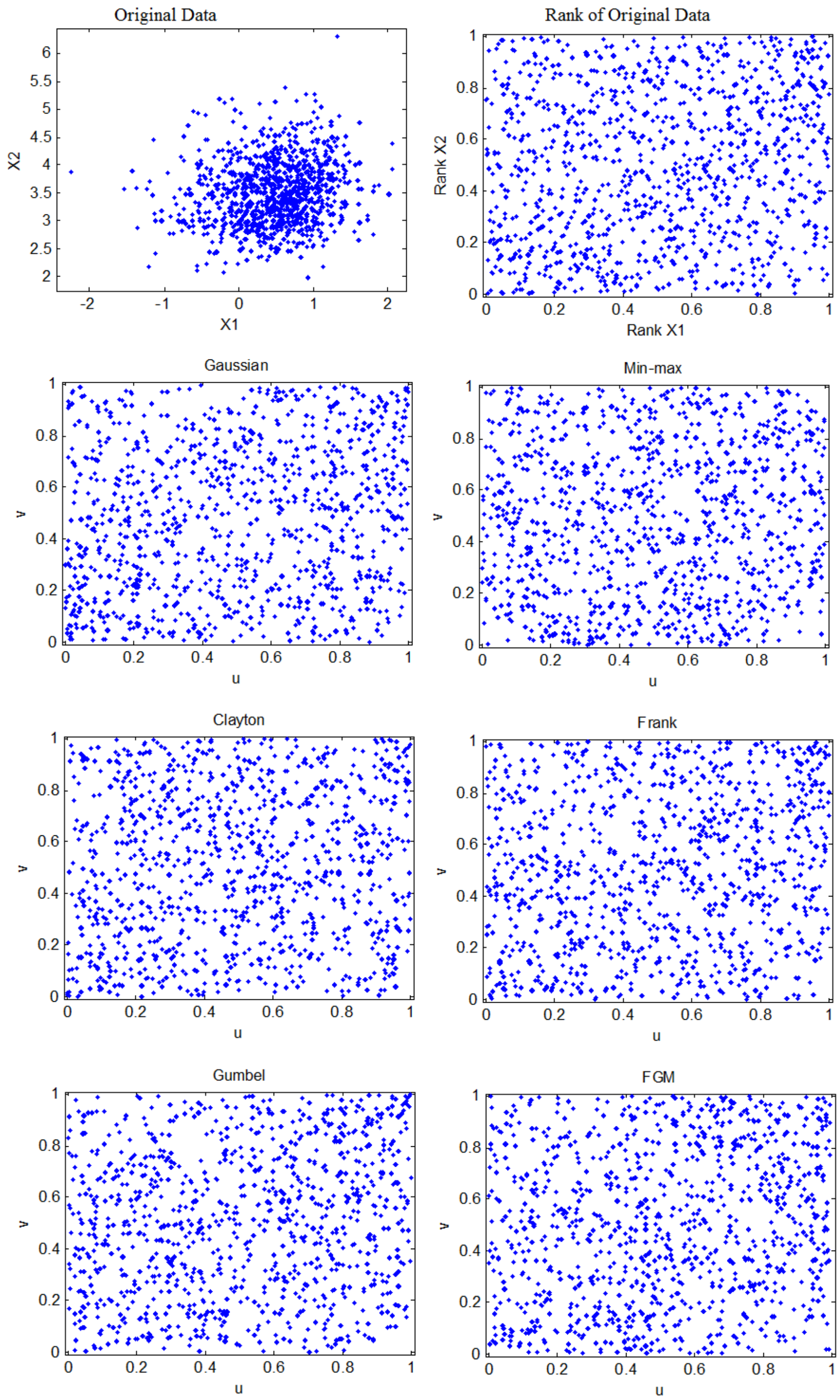

Figure 2. Scatter plots of original data, ranks of original data and copulas for $n = 10$

Figure 1 shows that min-max copula matches exactly the ranks of original data and we can easily visible min-max copula as the best-fit copula for $n=2$. For $n=10$, it is very difficult distinguish the best-fit copula among them in Figure 2 because of the ranks of the original data is very complicated and scattered. Also, it is very difficult to distinguish to the best-fit copula for $n=20, 100$, too

6. CONCLUSION

The study is intended to explore the dependency structure of the extreme order statistics which represents the minimum and maximum of n i.i.d. variables. The change in dependency is examined for different n values by means of simulation study and it is also aimed to find the copula family most suitable to explain dependency. Rank based methods is used for model selection and also it is tried to explain why rank-based methods are used. According to the results, for small n values ($n=2, 5$) only min-max copula is found suitable for dataset and when we examined scatter plots for $n=2$, we can easily distinguish the best-fit copula as min-max copula. For $n=10, 20, 100$, all copula families are founded suitable for data set but the best-fit copula is determined as min-max copula by means of AIC and BIC. Here, Clayton copula is the worst copula among the other copulas for $n=10$. However, Clayton copula is found better than Gaussian, Frank, Gumbel, FGM copulas for $n=20, 100$. For $n=200$ no copula families has been founded suitable for data set. Eventually, according the simulation results, min-max copula is the best one for small n values, however, it is not very useful for large n values.

REFERENCES

[1] Li X., Fang R., "Ordering properties of order statistics from random variables of Archimedean copulas with applications", ***Journal of Multivariate Analysis***, 133: 304-320, (2015).

[2] Arnold B.C., Balakrishnan N., Nagaraja H.N., "***A First Course in Order Statistics***", Vol. 54, John Wiley&Sons, New York, (1992).

[3] Balakrishnan N., Rao C. R., "***Handbook of Statistics 16: Order Statistics: Theory and Methods***", Elsevier, New York, (1998a).

[4] Balakrishnan N., Rao C. R., "***Handbook of Statistics 17: Order Statistics: Applications***", Vol. 42, Elsevier, New York, (1998b).

[5] David H. A., Nagaraja H. N., "***Order Statistics***", Wiley&Sons, Third edition, New Jersey, (2003).

[6] Schmitz V., "Revealing the dependence structure between $X_{(1)}$ and $X_{(n)}$", ***Journal of Statistical Planning Inference***, 123(1): 41–47, (2004).

[7] Li X., Li Z., "Proof of a conjecture on Spearman's ρ and Kendall's τ for sample minimum and maximum", ***Journal of Statistical Planning Inference***, 137(1): 359–361, (2007).

[8] Chen YP., "A note on the relationship between Spearman's ρ and Kendall's τ for extreme order statistics", ***Journal of Statistical Planning Inference***, 137(7): 2165–2171, (2007).

[9] Ghalibaf M. B., "Dependent structure of extreme order statistics", ***Journal of Statistical Computation and Simulation***, 86: 2846-2855, (2016).

[10] Li X., Fang R., "Ordering properties of order statistics from random variables of Archimedean copulas with applications", ***Journal of Multivariate Analysis***, 133: 304-320, (2015).

[11] Mesfioui M., Kayid M., Izadkhah S., "Stochastic comparisons of order statistics from heterogeneous random variables with Archimedean copula", ***Metrika***, 80: 749-766, (2017).

[12] Nelsen R.B., "***An Introduction to Copulas, in: Lecture Notes in Statistics***" Vol. 139, Springer,New York, (1999).

[13] Embrechts P., Lindskog F., McNeil A., "***Modeling dependence with copulas and applications to risk management***", In: S.T. Rachev, editor, Handbook of Heavy Tailed Distribution in Finance, JAI Press: Handbooks in Finance, 357-360, (2003).

[14] Genest C., Favre A.C., "Everything you always wanted to know about copula modeling but were afraid to ask", ***Journal of Hydrologic Engineering***, 12(4): 347-368, (2007).

[15] Li D.Q., Tang X.S., Phoon K.K. and Chen Y.F., "Bivariate simulation using copula and its application to probabilistic pile settlement analysis", ***International Journal Numerical Analytical Methods Geomechanics***, 37: 597-617, (2013).

[16] Genest C., Rivest L. P.," Statistical inference procedures for bivariate Archimedean copulas", ***Journal of the American Statistical Association***, 88(3): 1034-1043 (1993).

[17] Akaike H., "A new look at the statistical model identification", ***IEEE Transactions on Automatic Control***, 19(6): 716-723, (1974).

[18] Schwarz G., "Estimating the dimension of a model", ***The annals of Statistics***, 6(2): 461-464, (1978).

[19] Avérous, J.,Genest, C., Kochar, S.C. "On the dependence structure of order statistics", ***Journal of Multivariate Analysis***, 94(1): 159–171, (2005).