Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 52 (1) (2023), 209 – 228
DOI : 10.15672/hujms.1031959

Research Article

A consistent statistical test based on bivariate random samples

Aysegul Erem*[1], Ismihan Bayramoglu[2]

[1] *Department of Basic Sciences and Humanities, Faculty of Arts and Sciences, Cyprus International University, Lefkosa, Mersin 10, Turkey*

[2] *Department of Mathematics, Izmir University of Economics, Izmir, Turkey*

Abstract

We propose a consistent test for testing the distribution of bivariate random samples. The probability of type I, type II errors and probability of making no decisions under null and alternative hypotheses are obtained based on copula functions. The consistency of the proposed test is discussed under some null and alternative hypotheses. An unbiased, consistent estimator is proposed for probability of making no decision. Moreover, a simulation study is performed for showing the consistency of the proposed test for some well-known copulas such as independent, Clayton, Gumbel, Frank and Farlie-Gumbel-Morgenstern.

Mathematics Subject Classification (2020). 62G30, 62H15

Keywords. Bivariate two sample problem, bivariate order statistics, copula, hypothesis test, consistent tests

1. Introduction

One of the classical and important problems of mathematical statistics is two sample problem which has many applications in different areas such as biology, medicine, economics, ecology and engineering. This problem aroused the interest of many researchers and there is a tremendous number of research works dealing with two sample problems in the univariate cases. With the developments of machine learning methods in classifications of statistical data, this problem becomes attractive to many researchers. The classical statement of the problem considers two training samples from two populations and a control sample which must be classified to one of these populations. The classical theory of hypothesis testing offers many interesting criteria based on distances between distribution functions, such as Kolmogorov-Smirnov, Cramer-Von Mises, Chi-square etc. Using these criteria the researchers accept or reject the hypothesis asserting that the control sample belongs to one of two populations and calculating the probability of miss classification expressed in probabilities of type I and type II errors. The consistency and efficiency of criteria are subject to the large sample sizes are also important.

In machine learning, classification based on a rejection option is widely preferred. In rejection option technique, if the patterns are most likely misclassified, then they are

*Corresponding Author.
Email addresses: aysegulerem@gmail.com (A. Erem), ismihan.bayramoglu@izmirekonomi.edu.tr (I. Bayramoglu)
Received: 03.12.2021; Accepted: 28.07.2022

rejected (i.e., they are not classified). Then these unclassified patterns can be treated by more advanced classification techniques or they can be left to human operators for classification. As a result the researcher is averted from cost of making unreliable decisions. Fumera et al. [12] proposed multiple rejection thresholds related to the data classes. Herbei and Wegkamp [13] studied on a binary classification method with a rejection option which is constructed on plug-in rules and empirical risk minimizers. For more studies about the rejection option we also refer to [17, 18, 21, 27]

In a univariate case Bairamov and Petunin [2] proposed a consistent nonparametric test based on univariate training samples which are relied on the order statistics (OSs) and use of different test procedure. The null hypothesis assumes that the control sample belongs to one of the populations. It may be accepted or rejected if the control sample belongs to one of two intervals determined by OSs of training samples. If the control sample doesn't belong to the first or second interval, then no decision is taken and further investigations are required. The test is said to be consistent if probability of type I and type II errors both become infinitely small and if the sample size increases, the probability of making no decision is not one.

Computer-aided diagnostic systems are very important in medical sciences, when the diseases are diagnosed on the base of training samples of complex of symptoms of two groups of patients diagnosed with the disease A and B, respectively. The control sample is the complexes of symptoms of new patient who is determined to suffer with one of the diseases A or B and must be diagnosed to A or B by computer-aided diagnosis system. For example, A may be diabetes with normal blood sugar level (Diabetes Insipidus) and B may be diabetes with high blood sugar level (Diabetes Mellitus). However, in some cases it can be difficult to diagnose the case by a computer-aided diagnostic system. In such cases, the cases are examined by the expert as an advanced stage. Thus, the workload of the experts is reduced and the loss of time is minimized. If the probability of type I and type II errors are small and the probability of making no decision is large, the patient should be investigated in more detail until a correct diagnosis is made. For studies about the applications of rejection option in medicine, we refer to [7, 14, 19].

There are many interesting recent papers in statistical literature related to equality of two probability distributions, see, e.g., [1, 23–25]. Jiménez-Gamero et al. [15] proposed a test for testing equality of distribution of two samples based on empirical characteristic functions. Similarly, Alba-Fernández et al. [1] studied on a test statistic which relies on probability generating function. For more recent studies about univariate two-sample problems we also refer to [20].

In the case of bivariate observations, nonparametric two sample problem has not been studied much. These type of tests naturally involve the bivariate OSs and concomitants of order statistics. The theory of bivariate OSs is closely related to bivariate binomial distributions. Eryilmaz and Bairamov [10] studied the distribution of the bivariate training sample ranks of OSs and its concomitants. For some recent work on this topic we can refer to [4, 16]. There are some works dealing with the bivariate OSs and concomitants. Kemalbay and Bayramoglu [16] considered the joint distribution of ranks of OSs based on bivariate random samples. Stoimenova and Balakrishnan [24] also dealt with the two sample problem constructing a consistent test using precedence and exceedance statistics. Afterwards, Erem and Bayramoglu [9] obtained the exact and asymptotic distributions of exceedance statistics based on bivariate random sequences. Recently, Erem [8] proposed bivariate two sample test based on exceedance statistics for testing equality of two copula functions. For more results on the studies investigating the equality of two copula functions we refer to [6, 22]. Also, Susam and Ucer [26] studied the independency of Archimedean copulas based on Bernstein estimate of Kendall distribution function.

In this paper we consider two sample problem for bivariate observations. Similar to Bairamov and Petunin [2] considering univariate case, we construct statistical tests based on bivariate OSs for bivariate random samples. We divide a possible set of training samples into three mutually exclusive sets. If the control sample belongs to the first set we accept the null hypothesis. If it belongs to the second set we reject it and if it belongs to third set we do not make any decision and further investigations are made. The probabilities of type I and type II errors and also the probability of making no decision are calculated. The unbiased and consistent estimators for probability of not making a decision are constructed. The test proposed in this work can be used in many applications, in particular may be successful in the case of small sample observations.

In this paper we propose a test statistic for testing the equality of two copula functions based on bivariate OSs of training samples. In Section 2 problem statement and test procedure are introduced. Under some cases, the consistency of the test is discussed based on probability of type I, type II errors and probability of making no decision. In Section 3 an unbiased and consistent estimator for probability of making no decision is proposed. Finally in Section 4, a simulation study is performed for probability of type I, type II errors and probability of making no decision under some well-known copulas such as independent, Clayton, Frank, Gumbel and Farlie-Gumbel-Morgenstern (FGM) copulas.

2. Problem statement and the test procedure

Let $Z_1 = \left\{\left(X_k^{(1)}, Y_k^{(1)}\right), k = 1, 2, ..., n\right\}$ be a sequence of independent random variables with joint cumulative distribution function (cdf) $F(x, y) = C_1(F_X(x), F_Y(y))$, where $C_1(u, v)$, $(u, v) \in [0, 1]^2$ is a connecting copula and $F_X(x), F_Y(y)$ are the marginal cdf' s of X and Y, respectively. Furthermore, let $Z_2 = \left\{\left(X_k^{(2)}, Y_k^{(2)}\right), k = 1, 2, ..., n\right\}$ be another sequence of independent random variables with joint cdf $G(x, y) = C_2(F_X(x), F_Y(y))$, where $C_2(u, v)$, $(u, v) \in [0, 1]^2$ is a connecting copula and $F_X(x), F_Y(y)$ are the marginal cdf' s of X and Y, respectively. Let $f(x, y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$, $g(x, y) = \frac{\partial^2 G(x,y)}{\partial x \partial y}$, $f_X(x) = \frac{dF_X(x)}{dx}$ and $f_Y(y) = \frac{dF_Y(y)}{dy}$. We assume that Z_1 and Z_2 are independent and we call them the training samples from populations with joint cdfs F and G, respectively.

Let $Z = \{(X_1, Y_1), ..., (X_n, Y_n)\}$ be a control sample with joint distribution function $H(x, y) = C(F_X(x), F_Y(y))$, where $C(u, v)$, $(u, v) \in [0, 1]^2$ is a connecting copula. Recall Z_1, Z_2 are training samples and Z is a control sample. Consider the null hypothesis $H_0 : H(x, y) = F(x, y)$ and the alternative hypothesis $H_1 : H(x, y) = G(x, y)$. To test H_0 against alternative H_1 we construct a test based on bivariate OSs of the samples Z_1 and Z_2.

The (i, j)th bivariate OSs of Z_1 are defined as $(X_{i:n}^{(1)}, Y_{j:n}^{(1)})$, where $1 \leq i < j \leq n$, $X_{1:n}^{(1)} \leq X_{2:n}^{(1)} \leq \cdots \leq X_{n:n}^{(1)}$ and $Y_{1:n}^{(1)} \leq Y_{2:n}^{(1)} \leq \cdots \leq Y_{n:n}^{(1)}$ are the OSs of $\{X_k^{(1)}, k = 1, 2, ..., n\}$ and $\{Y_k^{(1)}, k = 1, 2, ..., n\}$, respectively. Similarly, the bivariate (i, j)th order statistics of Z_2 is $(X_{i:n}^{(2)}, Y_{j:n}^{(2)})$, where $X_{1:n}^{(2)} \leq X_{2:n}^{(2)} \leq \cdots \leq X_{n:n}^{(2)}$ and $Y_{1:n}^{(2)} \leq Y_{2:n}^{(2)} \leq \cdots \leq Y_{n:n}^{(2)}$ are the OSs of $\{X_k^{(2)}, k = 1, 2, ..., n\}$ and $\{Y_k^{(2)}, k = 1, 2, ..., n\}$, respectively.

We assume that $X_{1:n}^{(1)} \leq X_{1:n}^{(2)}$ and $X_{n:n}^{(1)} \leq X_{n:n}^{(2)}$. Similarly $Y_{1:n}^{(1)} \leq Y_{1:n}^{(2)}$ and $Y_{n:n}^{(1)} \leq Y_{n:n}^{(2)}$.

2.1. The test

Let $E_1 \equiv E_1(X, Y) = \left(-\infty, X_{1:n}^{(2)}\right) \times \left(-\infty, Y_{1:n}^{(2)}\right)$ and $E_2 \equiv E_2(X, Y) = \left(X_{n:n}^{(1)}, \infty\right) \times \left(Y_{n:n}^{(1)}, \infty\right)$. The test $T(n, r)$ $(1 < r < n)$ for testing H_0 against H_1 is defined as follows:

(1) Accept the hypothesis H_0 if at least r of the sample values of Z are in E_1.
(2) Accept H_1 if at least r of the sample values of Z are in E_2.
(3) In all other cases no decision is made about H_0 and H_1.

Since the sets E_1 and E_2 are disjoint, the alternatives 1 and 2 can not be observed at the same time. The graphical illustration of E_1 and E_2 is given in Figure 1.

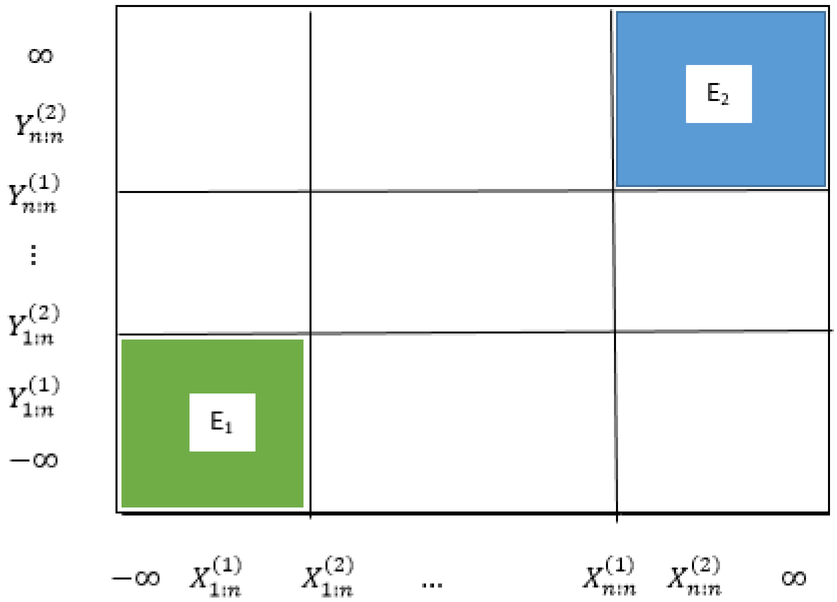

Figure 1. Graphical illustration of critical set.

Remark. To apply this test we need the following conditions:

$$X^{(1)}_{1:n} \le X^{(2)}_{1:n} \text{ and } X^{(1)}_{n:n} \le X^{(2)}_{n:n}, \text{ similarly } Y^{(1)}_{1:n} \le Y^{(2)}_{1:n} \text{ and } Y^{(1)}_{n:n} \le Y^{(2)}_{n:n}. \tag{2.1}$$

One of the referees rightly observes that the condition in Equation (2.1) is restrictive and will put strong relationship between two distribution functions F and G. This condition is satisfied in some situations if the training samples are not separated from each other, they are mixed. The verification of this condition is not difficult, because it is put on the data and required to check the training samples. The test can be applied and will be consistent if this condition is satisfied. In many practical applications the data from different samples are mixed, not separated. Alternatively, if

$$X^{(2)}_{1:n} \le X^{(1)}_{1:n} \text{ and } X^{(2)}_{n:n} \le X^{(1)}_{n:n}, \text{ similarly } Y^{(2)}_{1:n} \le Y^{(1)}_{1:n} \text{ and } Y^{(2)}_{n:n} \le Y^{(1)}_{n:n},$$

then the test procedure will be applied by changing places of the training samples Z_1 and Z_2. In the case when the training samples are separated, i.e.

$$X^{(1)}_{n:n} \le X^{(2)}_{1:n} \text{ and } Y^{(1)}_{n:n} \le Y^{(2)}_{1:n},$$

and in other cases not satisfying (2.1) different test procedures can be investigated and applied.

Definition 2.1. The test $T(n,r)$ based on bivariate training samples Z_1, Z_2 and control sample Z, is called consistent if probability of making no decision (ND) given that the hypothesis $H_i, i = 0,1$ ($P(ND \mid H_i)$) is true, is less than 1, i.e. $P(ND \mid H_i) < 1, i = 0,1$ and the probability of type I and type II errors tend to zero when the sample size n increases to infinity.

2.2. The probability of type I error

The probability of type I error α_n of the test $T(n,r)$ can be calculated as follows:

$$\begin{aligned}
\alpha_n &= P\left\{H_1 \mid H_0\right\}\\
&= P\Big\{\text{at least } r \text{ of the sample values in } Z \text{ are in } \left(X^{(1)}_{n:n},\infty\right)\\
&\qquad \times\left(Y^{(1)}_{n:n},\infty\right) \mid H(x,y)=F\left(x,y\right)\Big\}\\
&= P\left\{X_{n-r+1:n} > X^{(1)}_{n:n}, Y_{n-r+1:n} > Y^{(1)}_{n:n} \mid H(x,y)=F\left(x,y\right)\right\}\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[F(t,s)\right]^n f_{X_{n-r+1:n},Y_{n-r+1:n}}(t,s)dtds, \qquad (2.2)
\end{aligned}$$

where $f_{X_{n-r+1:n},Y_{n-r+1:n}}$ is probability density function (pdf) of $X^{(1)}_{n-r+1:n}$ and $Y^{(1)}_{n-r+1:n}$ in training sample Z_1. The joint pdf $f_{X_{n-r+1:n},Y_{n-r+1:n}}$ of bivariate OSs $X^{(1)}_{n-r+1:n}$ and $Y^{(1)}_{n-r+1:n}$ is studied in [3] and [16] has the following form:

$$\begin{aligned}
f_{X_{n-r+1:n},Y_{n-r+1:n}}(t,s) &= \sum_{t_1=a_1}^{a_2} l_1\left[F\left(t,s\right)\right]^{t_1}\left[\left(F_X\left(t\right)-F_1\left(t,s\right)\right)\left(F_Y\left(s\right)-F_1\left(t,s\right)\right)\right]^{n-r-t_1}\\
&\quad\times\left[\overline{F}\left(t,s\right)\right]^{2r-n-1+t_1} f\left(t,s\right)+\sum_{t_4=d_1}^{d_2}\sum_{t_2=c_1}^{c_2}\sum_{t_1=b_1}^{b_2} l_2\left[F\left(t,s\right)\right]^{t_1}\\
&\quad\times\left[\left(F_X\left(t\right)-F\left(t,s\right)\right)\right]^{n-r-t_1-t_2}\left[\left(F_Y\left(s\right)-F\left(t,s\right)\right)\right]^{n-r-t_1-t_4}\\
&\quad\times\left[\overline{F}\left(t,s\right)\right]^{2r-n+t_1+t_2+t_4-2}\left[F^{\cdot,1}\left(t,s\right)\right]^{t_2}\left[f_Y\left(s\right)-F^{\cdot,1}\left(t,s\right)\right]^{1-t_2}\\
&\quad\times\left[F^{1,\cdot}\left(t,s\right)\right]^{t_4}\left[f_X\left(t\right)-F^{1,\cdot}\left(t,s\right)\right]^{1-t_4}, \qquad (2.3)
\end{aligned}$$

where $a_1 = \max\left(0,n-2r+1\right)$, $a_2 = n-r$, $b_1 = \max\left(0,n-2r+2-t_2-t_4\right)$, $b_2 = \min\left(n-r-t_2,n-r-t_4\right)$, $c_1 = \max\left(0,2-r\right)$, $c_2 = \min\left(1,n-r\right)$, $d_1 = \max\left(0,2-r\right)$, $d_2 = \min\left(1,n-r\right)$,

$$F^{1,\cdot}\left(t,s\right) = \frac{\partial F_1\left(t,s\right)}{\partial t},$$
$$F^{\cdot,1}\left(t,s\right) = \frac{\partial F_1\left(t,s\right)}{\partial s},$$

and the constants l_1 and l_2 are

$$l_1 = \frac{n!}{t_1!\left[\left(n-r-t_1\right)!\right]^2\left(2r-n+t_1-1\right)!},$$
$$l_2 = \frac{n!}{t_1!\left(n-r-t_1-t_2\right)!\left(n-r-t_1-t_4\right)!\left(2r-n-2+t_1+t_2+t_4\right)!}.$$

2.3. The probability of type II error

The probability of Type II error β_n of the test $T(n,r)$ can be calculated as follows:

$$\begin{aligned}
\beta_n &= P\left\{H_0 \mid H_1\right\}\\
&= P\Big\{\text{at least } r \text{ of the sample values in } Z \text{ are in } \left(-\infty,X^{(2)}_{1:n}\right)\\
&\qquad \times\left(-\infty,Y^{(2)}_{1:n}\right) \mid H(x,y)=G\left(x,y\right)\Big\}\\
&= P\left\{X_{r:n} < X^{(2)}_{1:n}, Y_{n-1:n} < Y^{(2)}_{1:n} \mid H(x,y)=G\left(x,y\right)\right\}\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\overline{G}(t,s)\right]^n g_{X_{r:n},Y_{r:n}}\left(t,s\right)dtds, \qquad (2.4)
\end{aligned}$$

where $g_{X_{r:n},Y_{r:n}}(t,s)$ is pdf of $X_{r:n}^{(2)}$ and $Y_{r:n}^{(2)}$ in training sample Z_2 and $\overline{G}(t,s) = 1 - F_X(t) - F_Y(s) + G(t,s)$.

2.4. Consistency of the test

By using Equations (2.2) and (2.4) we have the following theorems.

Theorem 2.2. *The test $T(n,n-1)$ is consistent under the following hypotheses*

$$H_0 : C(u,v) = C_1(u,v) = uv$$
$$H_1 : C(u,v) = C_2(u,v).$$

Proof. See Appendix A. □

In Figure 2, the plot of α_n is provided for $C_1(u,v) = uv$.

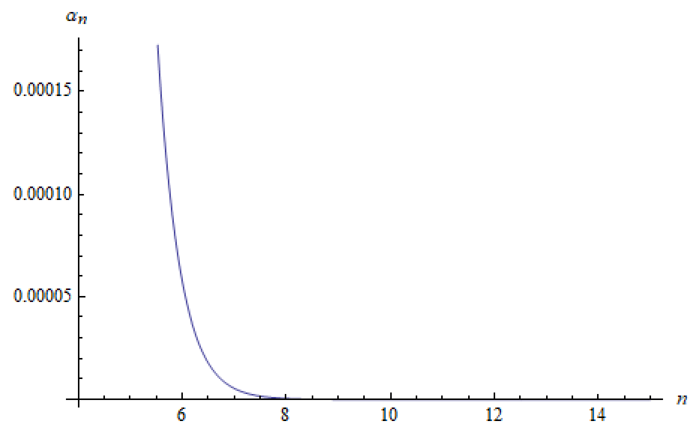

Figure 2. The plot of α_n for $C_1(u,v) = uv$.

Theorem 2.3. *The test $T(n,n-1)$ is consistent under $H_0 : C(u,v) = C_1(u,v)$ against $H_1 : C(u,v) = C_2(u,v) = uv$.*

Proof. See Appendix B. □

In Figure 3, the plot of β_n is provided for $C_2(u,v) = uv$.

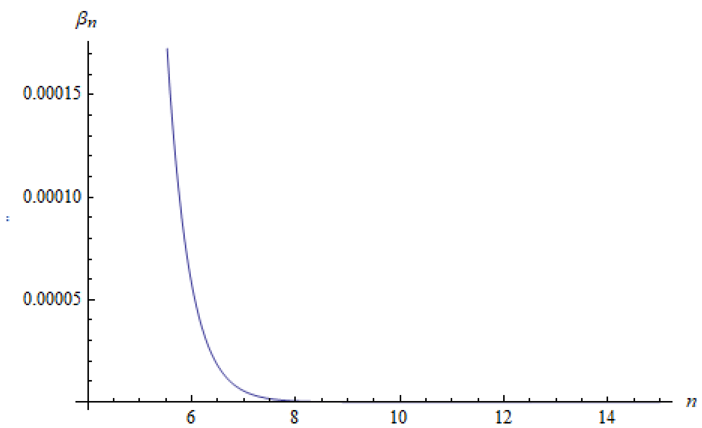

Figure 3. The plot of β_n for $C_2(u,v) = uv$.

3. An unbiased and consistent estimator of probability of making no decision

In this section, an unbiased and consistent estimator of probability of making no decisions under the null and alternative hypotheses ($P\{ND \mid H_0\}$ and $P\{ND \mid H_1\}$) are constructed. Here, ND denotes the event making no decision. By constructing the unbiased and consistent estimator, the investigator can make inference on probability of making no decision under the null hypothesis before using $T(n, n-1)$ test.

Firstly consider the unbiased and consistent estimator of $P\{ND \mid H_0\}$ under the following hypotheses:

$$H_0 : C(u,v) = C_1(u,v) = uv$$
$$H_1 : C(u,v) = C_2(u,v).$$

Let A denotes the event that at least r of the sample values of Z are in $\left(-\infty, X^{(2)}_{1:n}\right) \times \left(-\infty, Y^{(2)}_{1:n}\right)$ and B denotes the event at least r of the sample values of Z are in $\left(X^{(1)}_{n:n}, \infty\right) \times \left(Y^{(1)}_{n:n}, \infty\right)$. The events A and B are mutually exclusive events. Then by using Equation (A.4) in Appendix A, we have

$$P\{ND \mid H_0\} = 1 - \alpha_n - P\{A \mid H_0\}, \tag{3.1}$$

where

$$\alpha_n = \frac{n(n-1)^2(n+1)(n-2)!}{4^n A(n)} + \frac{(n^2-1)(n-1)!(n+1)!}{4^n (A(n))^2}.$$

Since $P\{A \mid H_0\}$ consists of a double integral which depends on unknown copula function $\widehat{C}_2(1-u, 1-v)$, an estimator for estimating the integral based on training samples Z_1 and Z_2 are constructed. The following sampling scheme is considered for bivariate training sample Z_1. We divide Z_1 sequentially into $l = \left[\frac{n}{k}\right]$ parts, such that

$$Z_1 = \left\{\left(X^{(1)}_1, Y^{(1)}_1\right), ..., \left(X^{(1)}_k, Y^{(1)}_k\right), \left(X^{(1)}_{k+1}, Y^{(1)}_{k+1}\right), ..., \left(X^{(1)}_{2k}, Y^{(1)}_{2k}\right), ..., \left(X^{(1)}_{k(l-1)}, Y^{(1)}_{k(l-1)}\right), ..., \left(X^{(1)}_{kl}, Y^{(1)}_{kl}\right)\right\},$$

and each sequential group is called as a subgroup of bivariate training sample of Z_1. Then, the OSs of i-th subgroup of training sample Z_1 with a sample size of k is

$$X^{(1)}_{i_{1:n}} \le X^{(1)}_{i_{2:n}} \le \cdots \le X^{(1)}_{i_{k:k}}$$
$$Y^{(1)}_{i_{1:n}} \le Y^{(1)}_{i_{2:n}} \le \cdots \le Y^{(1)}_{i_{k:k}}.$$

Define the event A_i as at least $k-1$ of $i-$th subgroup of Z_1 is in $\left(-\infty, X^{(2)}_{1:n}\right) \times \left(-\infty, Y^{(2)}_{1:n}\right)$. Then

$$\begin{aligned} P(A_i) &= P\left\{\text{at least } k-1 \text{ of } i-\text{th subgroup of } Z_1 \text{ is in } \left(-\infty, X^{(2)}_{1:n}\right) \times \left(-\infty, Y^{(2)}_{1:n}\right)\right\} \\ &= P\left\{X^{(1)}_{i_{k-1:k}} < X^{(2)}_{1:n}, Y^{(1)}_{i_{k-1:k}} < Y^{(2)}_{1:n}\right\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\overline{G}(t,s)\right]^n f_{X_{k-1:k}, Y_{k-1:k}}(t,s)\, dtds. \end{aligned} \tag{3.2}$$

Let define the binary random variable ξ_i as follows:

$$\xi_i = \begin{cases} 1 & \text{if event } A_i \text{ occurs,} \\ 0 & \text{if event } A_i \text{ does not occur,} \end{cases}$$

then $S_l = \xi_1 + \xi_2 + ... + \xi_l$ denotes the number of events A_i' s within l outcomes. It is clear that $P\{\xi_i = 1\} = P(A_i)$ and $P\{\xi_i = 0\} = 1 - P(A_i)$. From Equation (3.2) we can observe that $P(A_i)$ is equiprobable for $i = 1, 2, ..., l$. Let $h = \frac{S_l}{l}$ and $P(A_i) = p$. Then

$$E(h) = \frac{1}{l} E(S_l) = E(\xi_i) = p.$$

Therefore, $h = \frac{S_l}{l}$ is an unbiased estimator for $P\{A \mid H_0\}$ in Equation (3.1). Now consider the variance of statistic h

$$\begin{aligned} V(h) &= \frac{1}{l^2} V(S_l) \\ &= \frac{1}{l^2}\left(\sum_{i=1}^{l} V(\xi_i) + 2\sum_{i=1}^{l}\sum_{j=1}^{l} Cov(\xi_i, \xi_j)\right) \\ &= \frac{1}{l^2}\left(lp(1-p) + 2\binom{l}{2} Cov(\xi_i, \xi_j)\right) \\ &= \frac{p(1-p)}{l} + \frac{l-1}{l}\left[P(A_iA_j) - E(\xi_i)^2\right] \\ &= \frac{p(1-p)}{\left[\frac{n}{k}\right]} + \frac{\left[\frac{n}{k}\right]-1}{\left[\frac{n}{k}\right]}\left[P(A_iA_j) - p^2\right], \end{aligned}$$

where $P(A_iA_j) = p\{\xi_i = 1, \xi_j = 1\}$.

$$\begin{aligned} P(A_iA_j) &= P\left\{X^{(1)}_{i_{k-1:k}} < X^{(2)}_{1:n}, Y^{(1)}_{i_{k-1:k}} < Y^{(2)}_{1:n}, X^{(1)}_{j_{k-1:k}} < X^{(2)}_{1:n}, Y^{(1)}_{j_{k-1:k}} < Y^{(2)}_{1:n}\right\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P\left\{X^{(1)}_{i_{k-1:k}} < t, Y^{(1)}_{i_{k-1:k}} < s, X^{(1)}_{j_{k-1:k}} < t, Y^{(1)}_{j_{k-1:k}} < s\right\} g_{X_{1:n},Y_{1:n}}(t,s)\,dtds \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P\left\{X^{(1)}_{i_{k-1:k}} < t, Y^{(1)}_{i_{k-1:k}} < s\right\} P\left\{X^{(1)}_{j_{k-1:k}} < t, Y^{(1)}_{j_{k-1:k}} < s\right\} g_{X_{1:n},Y_{1:n}}(t,s)\,dtds \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[F_{X_{k-1:k},Y_{k-1:k}}(t,s)\right]^2 g_{X_{1:n},Y_{1:n}}(t,s)\,dtds, \end{aligned} \tag{3.3}$$

where $F_{X_{k-1:k},Y_{k-1:n}}(t,s)$ is the joint cdf of bivariate OSs $X^{(1)}_{k-1:k}$ and $Y^{(1)}_{k-1:k}$ in i-th subgroup of training sample Z_1 with a sample size of k. Also $g_{X_{1:n},Y_{1:n}}(t,s)$ is pdf of $X^{(2)}_{1:n}$ and $Y^{(2)}_{1:n}$ in training sample Z_2. By probability integral transformation, Equation (3.3) can be written for $C_1(u,v) = uv$ as follows:

$$\begin{aligned} P(A_iA_j) = \int_0^1\int_0^1 [C_{k,k}(u,v)]^2 &\left\{n\left[\widehat{C}_2(1-u,1-v)\right]^{n-1} c_2(u,v)\right. \\ &\left.+n(n-1)\left[\widehat{C}_2(1-u,1-v)\right]^{n-2}\left(1 - C_2^{\cdot,1}(u,v)\right)\left(1 - C_2^{1,\cdot}(u,v)\right)\right\} dudv, \end{aligned}$$

where $\widehat{C}_2(1-u,1-v) = 1-u-v+C_2(u,v)$, $c_2(u,v) = \frac{\partial^2 C_2(u,v)}{\partial u \partial v}$ and

$$\begin{aligned} C_{k,k}(u,v) = \sum_{i=k-1}^{k}\sum_{j=k-1}^{k}\sum_{l=\max(0,i+j-k)}^{\min(i,j)} &\frac{k!}{l!\,(i-l)!\,(j-l)!\,(k-i-j+l)!}\left[C_1(u,v)\right]^l\left[u - C_1(u,v)\right]^{i-l} \\ &\times\left[v - C_1(u,v)\right]^{j-l}\left[1-u-v+C_1(u,v)\right]^{k-i-j+l}. \end{aligned}$$

Then it is obvious that $\lim_{n\to\infty} P(A_iA_j) = 0$. Therefore, $\lim_{n\to\infty} V(h) = 0$.

So an unbiased and consistent estimator of $P\{ND \mid H_0\}$ under the following null and alternative hypotheses

$$H_0 : C(u,v) = C_1(u,v) = uv$$
$$H_1 : C(u,v) = C_2(u,v)$$

is given by

$$t_0(n) = 1 - \left[\frac{n(n-1)^2(n+1)(n-2)!}{4^n A(n)} + \frac{(n^2-1)(n-1)!\,(n+1)!}{4^n (A(n))^2} + h\right]. \tag{3.4}$$

In a similar way, one can also construct an unbiased and consistent estimator for $P\{ND \mid H_1\}$. Define the event B_i as at least $k-1$ of $i-$th subgroup of Z_2 is in $\left(X^{(1)}_{1:n}, \infty\right)\times$

$\left(Y_{1:n}^{(1)}, \infty\right)$. Then an unbiased and consistent estimator of $P\{ND \mid H_1\}$ under the following null and alternative hypotheses

$$H_0 : C(u,v) = C_1(u,v)$$
$$H_1 : C(u,v) = C_2(u,v) = uv$$

is given by

$$t_1(n) = 1 - \left[\frac{n(n-1)^2 [(n-2)!]^2 [(n+1)!]^2}{[(2n)!]^2} + \frac{(n-1)(n-1)!n![(n+1)!]^2}{[(2n)!]^2} + h^*\right],$$

where h^* is proportion of total number of events B_i' s in l subgroup of Z_2. In this section a simulation study is performed to show the consistency of the test under some well-known copulas (Independent, Clayton, Frank, Gumbel and FGM) in R programme for different values of n, r and dependence parameters. The number of repetition is 5000.

Table 1. Probability of type I Error for $C_1(u,v) = uv$ and different values of n and r.

n	r	α	n	r	α	n	r	α
10	1	0.072	20	1	0.0462	40	1	0.0192
10	2	0.0078	20	2	0.0016	40	2	0.0002
10	3	0.0014	20	3	0.0004	50	1	0.0188
10	4	0.0002	30	1	0.029	50	2	0.0006
			30	2	0.0012			

Table 2. Probability of type I Error for $C_1(u,v)$ is FGM copula and different values of n, r and θ_2.

(n,r)	θ_2	α	(n,r)	θ_2	α	(n,r)	θ_2	α
(10, 1)	−1	0.0224	(20, 1)	−1	0.0084	(30, 2)	−1	0
(10, 1)	0.5	0.0916	(20, 1)	0.5	0.061	(30, 2)	0.5	0.002
(10, 1)	1	0.1266	(20, 1)	1	0.094	(30, 2)	1	0.0062
(10, 2)	−1	0.0002	(20, 2)	−1	0.0002	(30, 3)	−1	0
(10, 2)	0.5	0.0134	(20, 2)	0.5	0.0052	(30, 3)	0.5	0.0002
(10, 2)	1	0.0232	(20, 2)	1	0.0088	(30, 3)	1	0.0006
(10, 3)	−1	0.0002	(20, 3)	−1	0	(50, 1)	−1	0.001
(10, 3)	0.5	0.0028	(20, 3)	0.5	0.0004	(50, 1)	0.5	0.027
(10, 3)	1	0.005	(20, 3)	1	0.002	(50, 1)	1	0.0388
(10, 4)	−1	0	(30, 1)	−1	0.0028	(50, 2)	−1	0
(10, 4)	0.5	0.004	(30, 1)	0.5	0.0412	(50, 2)	0.5	0.0012
(10, 4)	1	0.006	(30, 1)	1	0.052	(50, 2)	1	0.0022

Table 3. Probability of type I Error for $C_1(u,v)$ is Clayton copula and different values of n, r and θ_1.

(n,r)	θ_1	α	(n,r)	θ_1	α	(n,r)	θ_1	α	(n,r)	θ_1	α
(10, 1)	−1	0	(20, 1)	−1	0	(30, 1)	−1	0	(50, 1)	−1	0
(10, 1)	1	0.124	(20, 1)	1	0.0782	(30, 1)	1	0.0618	(50, 1)	1	0.034
(10, 1)	2	0.166	(20, 1)	2	0.1064	(30, 1)	2	0.0774	(50, 1)	2	0.0578
(10, 2)	−1	0	(20, 2)	−1	0	(30, 2)	−1	0	(50, 2)	−1	0
(10, 2)	1	0.023	(20, 2)	1	0.0104	(30, 2)	1	0.0032	(50, 2)	1	0.0024
(10, 2)	2	0.0352	(20, 2)	2	0.018	(30, 2)	2	0.0116	(50, 2)	2	0.006
(10, 5)	−1	0	(20, 3)	−1	0	(30, 3)	−1	0	(50, 3)	−1	0
(10, 5)	1	0	(20, 3)	1	0.0016	(30, 3)	1	0.0012	(50, 3)	1	0
(10, 5)	2	0.0012	(20, 3)	2	0.0036	(30, 3)	2	0.0016	(50, 3)	2	0.0006

In Tables 1-3, probability of type I error are provided under the following hypotheses:

$$\begin{aligned} H_0 &: C(u,v) = C_1(u,v) = uv \\ H_1 &: C(u,v) = C_2(u,v), \end{aligned} \tag{3.5}$$

$$H_0 : C(u,v) = C_1(u,v) = uv\left(1+\theta_2(1-u)(1-v)\right), \quad -1 \le \theta_2 \le 1, \text{ (FGM copula)}$$
$$H_1 : C(u,v) = C_2(u,v), \tag{3.6}$$

$$H_0 : C(u,v) = C_1(u,v) = \max\left(\left[u^{-\theta_1}+v^{-\theta_1}-1\right]^{-\frac{1}{\theta_1}},0\right), \ \theta_1 \in [-1,\infty)\setminus\{0\}, \text{(Clayton copula)}$$
$$H_1 : C(u,v) = C_2(u,v), \tag{3.7}$$

respectively. From Tables 1-3, it is clear that probability of type I error decreases when n increases. For fixed values of n, while r increases probability of type I error decreases. Furthermore for fixed values of n and r, when the dependence parameters (θ_1 and θ_2) increases, probability of type I error also increases. Similar conclusions can be done from Tables 4-6 for probability of type II error.

Table 4. Probability of type II Error for $C_2(u,v) = uv$ and different values of n and r.

n	r	β	n	r	β	n	r	β
10	1	0.081	20	1	0.04	40	1	0.022
10	2	0.0086	20	2	0.0034	40	2	0.0012
10	3	0.001	20	3	0	50	1	0.0172
10	4	0	30	1	0.0282	50	2	0.0014
			30	2	0.002			

Table 5. Probability of type II Error for $C_2(u,v)$ is FGM copula and different values of n, r and θ_2.

(n,r)	θ_2	β	(n,r)	θ_2	β	(n,r)	θ_2	β	(n,r)	θ_2	β
$(10,1)$	-1	0.021	$(20,1)$	-1	0.009	$(30,1)$	-1	0.004	$(50,1)$	-1	0.01
$(10,1)$	0.5	0.0984	$(20,1)$	0.5	0.0504	$(30,1)$	0.5	0.4	$(50,1)$	0.5	0.037
$(10,1)$	1	0.132	$(20,1)$	1	0.084	$(30,1)$	1	0.06	$(50,1)$	1	0.045
$(10,2)$	-1	0.001	$(20,2)$	-1	0	$(30,2)$	-1	0	$(50,2)$	-1	0
$(10,2)$	0.5	0.015	$(20,2)$	0.5	0.004	$(30,2)$	0.5	0.0032	$(50,2)$	0.5	0.003
$(10,2)$	1	0.028	$(20,2)$	1	0.01	$(30,2)$	1	0.0594	$(50,2)$	1	0.007
$(10,3)$	-1	0	$(20,3)$	-1	0	$(30,3)$	-1	0	$(10,4)$	-1	0
$(10,3)$	0.5	0.003	$(20,3)$	0.5	0	$(30,3)$	0.5	0	$(10,4)$	0.5	0
$(10,3)$	1	0.008	$(20,3)$	1	0.002	$(30,3)$	1	0.001	$(10,4)$	1	0.001

Table 6. Probability of type II Error for $C_2(u,v)$ is Clayton Copula and different values of n, r and θ_1.

(n,r)	θ_1	β	(n,r)	θ_1	β	(n,r)	θ_1	β	(n,r)	θ_1	β
$(10,1)$	-1	0	$(20,1)$	-1	0	$(30,1)$	-1	0	$(50,1)$	-1	0
$(10,1)$	1	0.317	$(20,1)$	1	0.305	$(30,1)$	1	0.285	$(50,1)$	1	0.274
$(10,1)$	2	0.41	$(20,1)$	2	0.392	$(30,1)$	2	0.383	$(50,1)$	2	0.4
$(10,2)$	-1	0	$(20,2)$	-1	0	$(30,2)$	-1	0	$(50,2)$	-1	0
$(10,2)$	1	0.096	$(20,2)$	1	0.087	$(30,2)$	1	0.081	$(50,2)$	1	0.074
$(10,2)$	2	0.142	$(20,2)$	2	0.182	$(30,2)$	2	0.155	$(50,2)$	2	0.144
$(10,3)$	-1	0	$(20,3)$	-1	0	$(30,3)$	-1	0	$(50,3)$	-1	0
$(10,3)$	1	0.001	$(20,3)$	1	0.081	$(30,3)$	1	0.024	$(50,3)$	1	0.0878
$(10,3)$	2	0.006	$(20,3)$	2	0.136	$(30,3)$	2	0.057	$(50,3)$	2	0.1458

In Tables 4-6 the simulation study is performed for probability of type II error under the following hypotheses, respectively:

$$H_0 : C(u,v) = C_1(u,v)$$
$$H_1 : C(u,v) = C_2(u,v) = uv,$$

$H_0 : C(u,v) = C_1(u,v)$

$H_1 : C(u,v) = C_2(u,v) = uv\left(1+\theta_2(1-u)(1-v)\right), \quad -1 \le \theta_2 \le 1,$ (FGM copula),

and

$H_0 : C(u,v) = C_1(u,v)$

$H_1 : C(u,v) = C_2(u,v) = \max\left(\left[u^{-\theta_1}+v^{-\theta_1}-1\right]^{-\frac{1}{\theta_1}}, 0\right), \ \theta_1 \in [-1,\infty)\setminus\{0\},$ (Clayton copula),

Table 7. $P(ND \mid H_0)$ under hypothesis (3.8) for some values of n, r and θ_1.

(n,r)	θ_1	$P(ND \mid H_0)$	(n,r)	θ_1	$P(ND \mid H_0)$
$(10,1)$	4	0.457	$(15,1)$	4	0.4978
$(10,1)$	5	0.437	$(15,1)$	5	0.4854
$(10,1)$	6	0.4144	$(15,1)$	6	0.4728
$(10,2)$	4	0.8282	$(15,2)$	4	0.8498
$(10,2)$	5	0.812	$(15,2)$	5	0.8414
$(10,2)$	6	0.8016	$(15,2)$	6	0.8306

Table 8. $P(ND \mid H_0)$ under hypothesis (3.9) for some values of n, r, θ_1 and θ_2.

(n,r)	(θ_1,θ_3)	$P(ND \mid H_0)$	(n,r)	(θ_1,θ_3)	$P(ND \mid H_0)$
$(10,1)$	$(1,5)$	0.56	$(15,1)$	$(1,5)$	0.5894
$(10,1)$	$(2,5)$	0.4526	$(15,1)$	$(2,5)$	0.5034
$(10,1)$	$(3,5)$	0.4026	$(15,1)$	$(3,5)$	0.4452
$(10,1)$	$(4,5)$	0.387	$(15,1)$	$(4,5)$	0.425
$(10,1)$	$(5,5)$	0.3458	$(15,1)$	$(5,5)$	0.4092
$(10,1)$	$(5,2)$	0.3998	$(15,1)$	$(5,2)$	0.457
$(10,1)$	$(5,3)$	0.368	$(15,1)$	$(5,3)$	0.4356
$(10,1)$	$(5,4)$	0.354	$(15,1)$	$(5,4)$	0.4088
$(10,1)$	$(5,6)$	0.3452	$(15,1)$	$(5,6)$	0.3936
$(10,2)$	$(1,5)$	0.8638	$(15,2)$	$(1,5)$	0.8852
$(10,2)$	$(2,5)$	0.8042	$(15,2)$	$(2,5)$	0.8824
$(10,2)$	$(3,5)$	0.7886	$(15,2)$	$(3,5)$	0.8014
$(10,2)$	$(4,5)$	0.7474	$(15,2)$	$(4,5)$	0.7822
$(10,2)$	$(5,5)$	0.7398	$(15,2)$	$(5,5)$	0.7696
$(10,2)$	$(5,2)$	0.7816	$(15,2)$	$(5,2)$	0.8188
$(10,2)$	$(5,3)$	0.7676	$(15,2)$	$(5,3)$	0.7918
$(10,2)$	$(5,4)$	0.7462	$(15,2)$	$(5,4)$	0.7766
$(10,2)$	$(5,6)$	0.7212	$(15,2)$	$(5,6)$	0.7588

Table 9. $P(ND \mid H_0)$ under hypothesis (3.10) for some values of n, r and θ_3.

(n,r)	θ_3	$P(ND \mid H_0)$	(n,r)	θ_3	$P(ND \mid H_0)$
$(10,1)$	2	0.458	$(15,1)$	2	0.4774
$(10,1)$	3	0.3458	$(15,1)$	3	0.3542
$(10,1)$	4	0.2822	$(15,1)$	4	0.3072
$(10,1)$	5	0.2652	$(15,1)$	5	0.2552
$(10,2)$	2	0.8362	$(15,2)$	2	0.8412
$(10,2)$	3	0.7674	$(15,2)$	3	0.772
$(10,2)$	4	0.728	$(15,2)$	4	0.738
$(10,2)$	5	0.6876	$(15,2)$	5	0.7116

In Tables 7-10, the simulation study is performed for probability of making no decision when the null hypothesis is true ($P(ND \mid H_0)$). We consider the following hypotheses, respectively:

$$H_0 : C(u,v) = C_1(u,v) = \max\left(\left[u^{-\theta_1}+v^{-\theta_1}-1\right]^{-\frac{1}{\theta_1}}, 0\right), \theta_1 \in [-1,\infty)\setminus\{0\}, \text{ (Clayton copula)}$$

$$H_1 : C(u,v) = C_2(u,v) = uv, \tag{3.8}$$

$$H_0 : C(u,v) = C_1(u,v) = \max\left(\left[u^{-\theta_1} + v^{-\theta_1} - 1\right]^{-\frac{1}{\theta_1}}, 0\right), \theta_1 \in [-1,\infty) \setminus \{0\}, \text{ (Clayton copula)},$$
$$H_1 : C(u,v) = C_2(u,v) = \exp\left\{-\left[(-\ln u)^{\theta_3} + (-\ln v)^{\theta_3}\right]^{1/\theta_3}\right\}, \theta_3 \in [1,\infty), \text{ (Gumbel copula)}, \tag{3.9}$$

$$H_0 : C(u,v) = C_1(u,v) = \exp\left\{-\left[(-\ln u)^{\theta_3} + (-\ln v)^{\theta_3}\right]^{1/\theta_3}\right\}, \theta_3 \in [1,\infty), \text{ (Gumbel copula)},$$
$$H_1 : C(u,v) = C_2(u,v) = uv, \tag{3.10}$$

$$H_0 : C(u,v) = C_1(u,v) = \exp\left\{-\left[(-\ln u)^{\theta_3} + (-\ln v)^{\theta_3}\right]^{1/\theta_3}\right\}, \ \theta_3 \in [1,\infty), \text{ (Gumbel copula)},$$
$$H_1 : C(u,v) = C_2(u,v) = -\frac{1}{\theta_4}\ln\left(1 + \frac{\left(e^{-\theta_4 u} - 1\right)\left(e^{-\theta_4 v} - 1\right)}{\left(e^{-\theta_4} - 1\right)}\right), \ \theta_4 \in (-\infty,\infty) \setminus \{0\}, \text{ (Frank copula)}. \tag{3.11}$$

Table 10. $P(ND \mid H_0)$ under hypothesis (3.11) for some values of n, r, θ_3 and θ_4.

(n,r)	(θ_3,θ_4)	$P(ND \mid H_0)$	(n,r)	(θ_3,θ_4)	$P(ND \mid H_0)$
$(10,1)$	$(3,5)$	0.3106	$(15,1)$	$(3,5)$	0.3332
$(10,1)$	$(4,5)$	0.2738	$(15,1)$	$(4,5)$	0.2892
$(10,1)$	$(5,5)$	0.2126	$(15,1)$	$(5,5)$	0.2382
$(10,1)$	$(5,2)$	0.2452	$(15,1)$	$(5,2)$	0.2482
$(10,1)$	$(5,3)$	0.2454	$(15,1)$	$(5,3)$	0.2572
$(10,1)$	$(5,4)$	0.2318	$(15,1)$	$(5,4)$	0.2638
$(10,2)$	$(3,5)$	0.7444	$(15,2)$	$(3,5)$	0.754
$(10,2)$	$(4,5)$	0.7008	$(15,2)$	$(4,5)$	0.7258
$(10,2)$	$(5,5)$	0.684	$(15,2)$	$(5,5)$	0.6834
$(10,2)$	$(5,2)$	0.6936	$(15,2)$	$(5,2)$	0.7032
$(10,2)$	$(5,3)$	0.6918	$(15,2)$	$(5,3)$	0.7008
$(10,2)$	$(5,4)$	0.6742	$(15,2)$	$(5,4)$	0.6838

From Tables 7-10, it can be observed that when n increases, $P(ND \mid H_0)$ increases. For fixed values of n and dependence parameters, when r increases $P(ND \mid H_0)$ increases. In Tables 7 and 9, when $\theta_1\,(\theta_3)$ increases $P(ND \mid H_0)$ decreases. In Table 8 for fixed values of n, r and $\theta_3\,(\theta_1)$, as $\theta_1\,(\theta_3)$ increases $P(ND \mid H_0)$ decreases. Furthermore, from Table 10, one can observe that for fixed values of n, r and $\theta_3\,(\theta_4)$, as $\theta_4\,(\theta_3)$ increases $P(ND \mid H_0)$ decreases.

In Tables 11-15 the results of simulation study are provided for probability of making no decision, when the alternative hypothesis is true ($P(ND \mid H_1)$). We consider the following hypotheses, respectively:

$$H_0 : C(u,v) = C_1(u,v) = uv,$$
$$H_1 : C(u,v) = C_2(u,v) = \max\left(\left[u^{-\theta_1} + v^{-\theta_1} - 1\right]^{-\frac{1}{\theta_1}}, 0\right), \ \theta_1 \in [-1,\infty) \setminus \{0\}, \text{ (Clayton copula)}, \tag{3.12}$$

$$H_0 : C(u,v) = C_1(u,v) = uv\left(1 + \theta_2(1-u)(1-v)\right), \ -1 \le \theta_2 \le 1 \text{ (FGM copula)},$$
$$H_1 : C(u,v) = C_2(u,v) = \max\left(\left[u^{-\theta_1} + v^{-\theta_1} - 1\right]^{-\frac{1}{\theta_1}}, 0\right), \theta_1 \in [-1,\infty) \setminus \{0\}, \text{ (Clayton copula)}, \tag{3.13}$$

$$H_0 : C(u,v) = C_1(u,v) = uv,$$
$$H_1 : C(u,v) = C_2(u,v) = \exp\left\{-\left[(-\ln u)^{\theta_3} + (-\ln v)^{\theta_3}\right]^{1/\theta_3}\right\}, \theta_3 \in [1,\infty), \text{ (Gumbel copula)}, \tag{3.14}$$

$$H_0 : C(u,v) = C_1(u,v) = -\frac{1}{\theta_4} \ln \left(1 + \frac{\left(e^{-\theta_4 u} - 1\right)\left(e^{-\theta_4 v} - 1\right)}{\left(e^{-\theta_4} - 1\right)} \right), \ \theta_4 \in (-\infty, \infty) \setminus \{0\}, \quad \text{(Frank copula)},$$

$$H_1 : C(u,v) = C_2(u,v) = \exp\left\{ - \left[(-\ln u)^{\theta_3} + (-\ln v)^{\theta_3} \right]^{1/\theta_3} \right\}, \ \theta_3 \in [1, \infty), \ \text{(Gumbel copula)}. \tag{3.15}$$

Table 11. $P(ND \mid H_1)$ under hypothesis (3.12) for n, r and θ_1.

(n,r)	θ_1	$P(ND \mid H_1)$	(n,r)	θ_1	$P(ND \mid H_1)$
$(10,1)$	3	0.4214	$(15,1)$	3	0.4312
$(10,1)$	4	0.3626	$(15,1)$	4	0.3836
$(10,1)$	5	0.3498	$(15,1)$	5	0.3802
$(10,1)$	6	0.342	$(15,1)$	6	0.3352
$(10,2)$	3	0.7766	$(15,2)$	3	0.7922
$(10,2)$	4	0.7604	$(15,2)$	4	0.7768
$(10,2)$	5	0.7416	$(15,2)$	5	0.761
$(10,2)$	6	0.7152	$(15,2)$	6	0.733

Table 12. $P(ND \mid H_1)$ under hypothesis (3.13) for some values of n, r, θ_1 and θ_2.

(n,r)	(θ_1,θ_2)	$P(ND \mid H_1)$	(n,r)	(θ_1,θ_2)	$P(ND \mid H_1)$
$(10,1)$	$(4,-1)$	0.3724	$(15,1)$	$(4,-1)$	0.411
$(10,1)$	$(4,1)$	0.3522	$(15,1)$	$(4,1)$	0.398
$(10,1)$	$(5,-1)$	0.3484	$(15,1)$	$(5,-1)$	0.3944
$(10,1)$	$(5,1)$	0.3346	$(15,1)$	$(5,1)$	0.3638
$(10,2)$	$(4,-1)$	0.7602	$(15,2)$	$(4,-1)$	0.7698
$(10,2)$	$(4,1)$	0.7338	$(15,2)$	$(4,1)$	0.7658
$(10,2)$	$(5,-1)$	0.746	$(15,2)$	$(5,-1)$	0.7512
$(10,2)$	$(5,1)$	0.721	$(15,2)$	$(5,1)$	0.7338

Table 13. $P(ND \mid H_1)$ under hypothesis (3.9) for some values of n, r, θ_1 and θ_2.

(n,r)	(θ_1,θ_3)	$P(ND \mid H_1)$	(n,r)	(θ_1,θ_3)	$P(ND \mid H_1)$
$(10,1)$	$(1,5)$	0.2872	$(15,1)$	$(1,5)$	0.2948
$(10,1)$	$(2,5)$	0.2808	$(15,1)$	$(2,5)$	0.3086
$(10,1)$	$(3,5)$	0.2768	$(15,1)$	$(3,5)$	0.306
$(10,1)$	$(4,5)$	0.267	$(15,1)$	$(4,5)$	0.2866
$(10,1)$	$(5,5)$	0.2558	$(15,1)$	$(5,5)$	0.301
$(10,1)$	$(5,2)$	0.4666	$(15,1)$	$(5,2)$	0.4956
$(10,1)$	$(5,3)$	0.3504	$(15,1)$	$(5,3)$	0.3758
$(10,1)$	$(5,4)$	0.3092	$(15,1)$	$(5,4)$	0.3154
$(10,1)$	$(5,6)$	0.2624	$(15,1)$	$(5,6)$	0.2896
$(10,2)$	$(1,5)$	0.7242	$(15,2)$	$(1,5)$	0.7436
$(10,2)$	$(2,5)$	0.7238	$(15,2)$	$(2,5)$	0.7524
$(10,2)$	$(3,5)$	0.7198	$(15,2)$	$(3,5)$	0.7202
$(10,2)$	$(4,5)$	0.7148	$(15,2)$	$(4,5)$	0.723
$(10,2)$	$(5,5)$	0.7004	$(15,2)$	$(5,5)$	0.7238
$(10,2)$	$(5,2)$	0.8402	$(15,2)$	$(5,2)$	0.8588
$(10,2)$	$(5,3)$	0.7662	$(15,2)$	$(5,3)$	0.7872
$(10,2)$	$(5,4)$	0.73	$(15,2)$	$(5,4)$	0.7502
$(10,2)$	$(5,6)$	0.695	$(15,2)$	$(5,6)$	0.6976

In Tables 11-15, one can observe that as n increases, $P(ND \mid H_1)$ increases for fixed values of r and dependence parameters. Furthermore for fixed values of r and the dependence parameters, as r increases $P(ND \mid H_1)$ decreases. In Tables 11 and 14, when

θ_1 (θ_3) increases $P(ND \mid H_1)$ decreases. In Table 12 for fixed values of θ_1 (θ_2), n and r when $\theta_2\,(\theta_1)$ increases $P(ND \mid H_1)$ decreases. In Table 13, for fixed values of n, r and θ_3 when θ_1 increases, $P(ND \mid H_1)$ decreases. Similarly for fixed values of n, r and θ_1 when θ_3 increases, $P(ND \mid H_1)$ decreases. Likewise, in Table 15 one can observe that for fixed values of θ_3 (θ_4), n and r, as θ_4 (θ_3) increases $P(ND \mid H_1)$ decreases.

Table 14. $P(ND \mid H_1)$ under hypothesis (3.10) for some values of n, r and θ_3.

(n,r)	θ_3	$P(ND \mid H_1)$	(n,r)	θ_3	$P(ND \mid H_1)$
$(10,1)$	2	0.4924	$(15,1)$	2	0.5182
$(10,1)$	3	0.3656	$(15,1)$	3	0.3982
$(10,1)$	4	0.3274	$(15,1)$	4	0.3308
$(10,1)$	5	0.2932	$(15,1)$	5	0.3164
$(10,2)$	2	0.863	$(15,2)$	2	0.875
$(10,2)$	3	0.7958	$(15,2)$	3	0.801
$(10,2)$	4	0.7566	$(15,2)$	4	0.7778
$(10,2)$	5	0.7316	$(15,2)$	5	0.7434

Table 15. $P(ND \mid H_1)$ under hypothesis (3.15) for some values of n, r, θ_3 and θ_4.

(n,r)	(θ_3,θ_4)	$P(ND \mid H_1)$	(n,r)	(θ_3,θ_4)	$P(ND \mid H_1)$
$(10,1)$	$(3,5)$	0.3488	$(15,1)$	$(3,5)$	0.4062
$(10,1)$	$(4,5)$	0.317	$(15,1)$	$(4,5)$	0.3332
$(10,1)$	$(5,5)$	0.2552	$(15,1)$	$(5,5)$	0.2874
$(10,1)$	$(5,2)$	0.294	$(15,1)$	$(5,2)$	0.3126
$(10,1)$	$(5,3)$	0.2786	$(15,1)$	$(5,3)$	0.3088
$(10,1)$	$(5,4)$	0.2764	$(15,1)$	$(5,4)$	0.301
$(10,2)$	$(3,5)$	0.7716	$(15,2)$	$(3,5)$	0.7882
$(10,2)$	$(4,5)$	0.7372	$(15,2)$	$(4,5)$	0.7536
$(10,2)$	$(5,5)$	0.7166	$(15,2)$	$(5,5)$	0.7378
$(10,2)$	$(5,2)$	0.7366	$(15,2)$	$(5,2)$	0.751
$(10,2)$	$(5,3)$	0.7254	$(15,2)$	$(5,3)$	0.7282
$(10,2)$	$(5,4)$	0.716	$(15,2)$	$(5,4)$	0.7348

4. Conclusion

In this paper, a test statistic $T(n,r)$ is proposed for testing the distribution of bivariate sample Z based on bivariate samples Z_1 and Z_2. The consistency of $T(n,n-1)$ test is discussed by considering the probability of type I, II errors and probability of making no decision under null and alternative hypotheses. These probabilities involve copulas of underlying distributions. Also unbiased and consistent estimators for probability of making no decision under the null/alternative hypotheses are proposed. This approach can help the investigators in prediction of probability of making no decision. If probability of making no decision is high, the test can not be performed or it can be repeated with different samples. Therefore, the researcher avoids making wrong decision. Furthermore, under particular hypotheses a simulation study is performed. In simulation study we observe that the probability of making no decision under the null/alternative hypotheses is considerably high for large values of n. Therefore, the use of $T(n,r)$ is preferable for small sample size.

Acknowledgment. The authors thank the anonymous reviewers and editor for their valuable comments and suggestions that led to improvements in the paper.

References

[1] M.V. Alba-Fernández, A. Batsidis, M.D. Jiménez-Gamero and P. Jodrá, *A class of tests for the two-sample problem for count data*, J. Comput. Appl. Math. **318**, 220-229, 2017.

[2] I.G. Bairamov and Y.I. Petunin, *Statistical tests based on training samples*, Cybernet. Systems Anal. **27** (3), 408-413, 1991.

[3] H.M. Barakat, *On moments of bivariate order statistics*, Ann. Inst. Statist. Math. **51** (2), 351-358, 1999.

[4] I. Bayramoglu and S. Eryilmaz, *Order statistics of dependent sequences consisting of two different sets of exchangeable variables*, J. Comput. Appl. Math. **286**, 1-6, 2015.

[5] I. Bayramoglu and G. Kemalbay, *Some novel discrete distributions under fourfold sampling schemes and conditional bivariate order statistics*, J. Comput. Appl. Math. **248**, 1-14, 2013.

[6] A. Bücher, H. Dette and S. Volgushev, *A test for archimedeanity in bivariate copula models*, J. Multivariate Anal. **110**, 121–132, 2012.

[7] A.R. da Rocha Neto, R. Sousa, G.D.A. Barreto and J.S. Cardoso, *Diagnostic of pathology on the vertebral column with embedded reject option*, in: Iberian Conference on Pattern Recognition and Image Analysis, 588-595, Springer, Berlin, Heidelberg, 2011.

[8] A. Erem, *Bivariate two sample test based on exceedance statistics*, Comm. Statist. Simulation Comput. **49** (9), 2020.

[9] A. Erem and I. Bayramoglu, *Exact and asymptotic distributions of exceedance statistics for bivariate random sequences*, Statist. Probab. Lett. **125**, 181-188, 2017.

[10] S. Eryılmaz and I.G. Bairamov, *On a new sample rank of an order statistics and its concomitant*, Statist. Probab. Lett. **63** (2), 123-131, 2003.

[11] V.A. Fernández, M.J. Gamero, and J.M. Garcia, *A test for the two-sample problem based on empirical characteristic functions*, Comput. Statist. Data Anal. **52** (7), 3730-3748, 2008.

[12] G. Fumera, F. Roli and G. Giacinto, *Reject option with multiple thresholds*, Pattern Recognit. **33** (12), 2099-2101, 2000.

[13] R. Herbei and M.H. Wegkamp, *Classification with reject option*, Canad. J. Statist. **34** (4), 709–721, 2006.

[14] W. Homenda, A. Jastrzebska and W. Pedrycz, *Unsupervised mode of rejection of foreign patterns*, Appl. Soft Comput. **57**, 615-626, 2017.

[15] M.D. Jiménez-Gamero, M.V. Alba-Fernández, P. Jodrá and I. Barranco-Chamorro, *Fast tests for the two-sample problem based on the empirical characteristic function*, Math. Comput. Simulation **137**, 390-410, 2017.

[16] G. Kemalbay and I. Bayramoglu, *Joint distribution of new sample rank of bivariate order statistics*, J. Appl. Stat. **42** (10), 2280-2289, 2015.

[17] D. Lin, L. Sun, K.A. Toh, J.B. Zhang and Z. Lin, *Biomedical image classification based on a cascade of an SVM with a reject option and subspace analysis*, Comput. Biol. Med. **96**, 128-140, 2018.

[18] D.P. Mesquita, L.S. Rocha, J.P.P. Gomes and A.R.R. Neto, *Classification with reject option for software defect prediction*, Appl. Soft Comput. **49**, 1085-1093, 2016.

[19] R.A. Mohammadpour, S.M. Abedi, S. Bagheri and A. Ghaemian, *Fuzzy rule-based classification system for assessing coronary artery disease*, Comput. Math. Methods Med., Doi: 10.1155/2015/564867, 2015.

[20] T. Nishino and H. Murakami, *The generalized Cucconi test statistic for the two-sample problem*, J. Korean Statist. Soc. **48** (4), 593-612, 2019.

[21] I. Pillai, G. Fumera and F. Roli, *Multi-label classification with a reject option*, Pattern Recognit. **46** (8), 2256-2266, 2013.

[22] B. Rémillard and O. Scaillet, *Testing for equality between two copulas*, J. Multivariate Anal. **100** (3), 377-386, 2009.
[23] E. Stoimenova, *The power of exceedance-type tests under lehmann alternatives*, Comm. Statist. Theory Methods **40** (4), 731-744, 2011.
[24] E. Stoimenova and N. Balakrishnan, *A class of exceedance-type statistics for the two-sample problem*, J. Statist. Plann. Inference **141** (9), 3244-3255, 2011.
[25] E. Stoimenova and N. Balakrishnan, *Sidak-type tests for the two-sample problem based on precedence and exceedance statistics*, Statistics **51** (2), 247-264, 2017.
[26] S.O. Susam and B. Hudaverdi Ucer, *Testing independence for Archimedean copula based on Bernstein estimate of Kendall distribution function*, J. Stat. Comput. Simul. **88** (13), 2589-2599, 2018.
[27] D.M. Tax and R.P. Duin, *Growing a multi-class classifier with a reject option*, Pattern Recognit. Lett. **29** (10), 1565-1570, 2008.

Appendix A

Proof of Theorem 2.2. We have to show that the probabilities of type I and type II errors both tend to infinity as $n \to \infty$ and the probability of not making a decision under hypotheses H_0 and H_1 are less than 1. We use probability integral transformation $F_X(t) = u$, $F_Y(s) = v$ and $F(t,s) = C_1(F_X(t), F_Y(s))$, in Equation (2.3). It is clear that

$$|J| = \begin{vmatrix} \frac{\partial u}{\partial t} & \frac{\partial u}{\partial s} \\ \frac{\partial v}{\partial t} & \frac{\partial v}{\partial s} \end{vmatrix} = \begin{vmatrix} \frac{1}{f_X(F_X^{-1}(u))} & 0 \\ 0 & \frac{1}{f_Y(F_Y^{-1}(v))} \end{vmatrix} = \frac{1}{f_X(t) f_Y(s)}$$

and

$$\begin{aligned} F^{1,\cdot}(t,s) &= \frac{\partial F(t,s)}{\partial t} = \frac{\partial C_1(F_X(t), F_Y(s))}{\partial t} \\ &= \frac{\partial C_1(F_X(t), F_Y(s))}{\partial F_X(t)} \cdot \frac{dF_X(t)}{dt} \\ &= C_1^{1,\cdot}(F_X(t), F_Y(s)) f_X(t), \end{aligned} \tag{A.1}$$

where $C_1^{1,\cdot}(F_X(t), F_Y(s)) = \frac{\partial C_1(F_X(t), F_Y(s))}{\partial F_X(t)}$. Similarly,

$$F_1^{\cdot,1}(t,s) = C_1^{\cdot,1}(F_X(t), F_Y(s)) f_Y(s),$$

where $C_1^{\cdot,1}(F_X(t), F_Y(s)) = \frac{\partial C_1(F_X(t), F_Y(s))}{\partial F_Y(s)}$. Taking into account Equations (2.2) and (2.3), the probability of type I error can be written as follows:

$$\begin{aligned} \alpha_n = \int_0^1 \int_0^1 (C_1(u,v))^n \Bigg\{ & \sum_{t_1=a_1}^{a_2} l_1 [C_1(u,v)]^{t_1} [(u - C_1(u,v))(v - C_1(u,v))]^{n-r-t_1} \\ & \times \left[\widehat{C}_1(1-u, 1-v)\right]^{2r-n-1+t_1} c_1(u,v) + \sum_{t_4=d_1}^{d_2} \sum_{t_2=c_1}^{c_2} \sum_{t_1=b_1}^{b_2} l_2 [C_1(u,v)]^{t_1} [u - C_1(u,v)]^{n-r-t_1-t_2} \\ & \times [v - C_1(u,v)]^{n-r-t_1-t_4} \left[\widehat{C}_1(1-u,1-v)\right]^{2r-n+t_1+t_2+t_4-2}, \\ & \times \left[C_1^{\cdot,1}(u,v)\right]^{t_2} \left[1 - C_1^{\cdot,1}(u,v)\right]^{1-t_2} \left[C_1^{1,\cdot}(u,v)\right]^{t_4} \left[1 - C_1^{1,\cdot}(u,v)\right]^{1-t_4} \Bigg\} dudv, \end{aligned} \tag{A.2}$$

where $\widehat{C}_1(1-u, 1-v) = 1 - u - v + C_1(u,v)$ and $c_1(u,v) = \frac{\partial^2 C_1(u,v)}{\partial u \partial v}$. It is obvious that $a_1 = c_1 = d_1 = 0$ and $a_2 = c_2 = d_2 = 1$ for $r = n-1$. Then, the probability of type I error under $H_0 : C(u,v) = C_1(u,v) = uv$ is given by

$$\begin{aligned}\alpha_n &= \int_0^1 \int_0^1 (uv)^n \left\{ \left(n(n-1)^2 uv \left[(1-u)(1-v)\right]^{n-2} \right) \right.\\ &\left. + \sum_{t_4=0}^{1} \sum_{t_2=0}^{1} \sum_{t_1=\max(0,4-n-t_2-t_4)}^{\min(1-t_2,1-t_4)} l_2 uv \left[(1-u)(1-v)\right]^{n-2} \right\} dudv \\ &= \frac{n(n-1)^2 (n+1)!(n-2)!}{(2n)!} + \frac{(n^2-1)(n-1)!(n+1)!(n!)^2}{((2n)!)^2} \\ &= \frac{n(n-1)^2 (n+1)(n-2)!}{4^n A(n)} + \frac{(n^2-1)(n-1)!(n+1)!}{4^n (A(n))^2},\end{aligned}$$

where

$$A(n) = \prod_{k=1}^{n} (2n - (2k-1)).$$

For simplicity denote $I_n^{'} = \frac{n(n-1)^2(n+1)(n-2)!}{4^n A(n)}$ and $I_n^{''} = \frac{(n^2-1)(n-1)!(n+1)!}{4^n (A(n))^2}$. Then,

$$\begin{aligned}\lim_{n\to\infty} I_n^{'} &= \lim_{n\to\infty} \frac{n(n-1)^2(n+1)(n-2)!}{4^n A(n)} \\ &= \lim_{n\to\infty} \frac{n(n+1)(n-2)!}{A(n)} \cdot \lim_{n\to\infty} \frac{(n-1)^2}{4^n} \\ &= 0 \cdot \lim_{n\to\infty} \frac{(n-1)^2}{4^n} \\ &= 0.\end{aligned}$$

Similarly,

$$\begin{aligned}\lim_{n\to\infty} I_n^{''} &= \lim_{n\to\infty} \frac{(n^2-1)(n-1)!(n+1)!}{4^n A(n)} \\ &= \lim_{n\to\infty} \frac{(n^2-1)(n+1)(n-1)!\,n!}{4^n (A(n))^2} \\ &= \lim_{n\to\infty} \frac{(n^2-1)(n+1)}{4^n} \cdot \lim_{n\to\infty} \frac{(n-1)!\,n!}{(A(n))^2} \\ &= 0.\end{aligned}$$

It can be easily shown that $\lim_{n\to\infty} \alpha_n = 0$.

Now let us calculate the probability of no decision given that H_0 is true. Let A denotes the event that at least r of the sample values in Z are in $\left(-\infty, X_{1:n}^{(2)}\right) \times \left(-\infty, Y_{1:n}^{(2)}\right)$ and B denotes the event at least r of the sample values in Z are in $\left(X_{n:n}^{(1)}, \infty\right) \times \left(Y_{n:n}^{(1)}, \infty\right)$. The events A and B are mutually exclusive events. Then,

$$\begin{aligned}P\{ND \mid H_0\} &= 1 - P\{A \cup B \mid H_0\} \\ &= 1 - P\{A \mid H_0\} - P\{B \mid H_0\}. \end{aligned} \tag{A.3}$$

It is obvious that $P\{B \mid H_0\} = \alpha_n$, then,

$$P\{ND \mid H_0\} = 1 - \alpha_n - P\{A \mid H_0\}.$$

$$\begin{aligned}
P\{A \mid H_0\} &= P\left\{\text{at least } r \text{ of the sample values in } Z \text{ are in } \left(-\infty, X_{1:n}^{(2)}\right) \times \left(-\infty, Y_{1:n}^{(2)}\right) \mid H_0\right\} \\
&= P\left\{X_{r:n} < X_{1:n}^{(2)}, Y_{r:n} < Y_{1:n}^{(2)} \mid H_0\right\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\overline{G}(t,s)\right]^n f_{X_{r:n},Y_{r:n}}(t,s)\, dtds. \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\overline{G}(t,s)\right]^n \left(\sum_{t_1=n-3}^{n-2} \widetilde{l}_1 \left[F(t,s)\right]^{t_1} \left[\left(F_X(t) - F(t,s)\right)\left(F_Y(s) - F(t,s)\right)\right]^{n-2-t_1}\right. \\
&\quad \left. \times \left[\overline{F}(t,s)\right]^{3-n+t_1} f(t,s) + \sum_{t_4=0}^{1}\sum_{t_2=0}^{1}\sum_{t_1=\max(0,n-2-t_2-t_4)}^{\min(n-2-t_2,n-2-t_4)} \widetilde{l}_2 \left[F(t,s)\right]^{t_1}\right),
\end{aligned}$$

where

$$\widetilde{l}_1 = \frac{n!}{t_1!\left[(n-2-t_1)!\right]^2 (3-n+t_1)!}$$

and

$$\widetilde{l}_2 = \frac{n!}{t_1!\,(n-2-t_1-t_2)!\,(n-2-t_1-t_4)!\,(2-n+t_1+t_2+t_4)!}.$$

Similarly by using probability integral transformation we have

$$\begin{aligned}
P\{A \mid H_0\} = \int_0^1\int_0^1 &\left(\widehat{C}_2(1-u,1-v)\right)^n \\
&\times \left\{\sum_{t_1=n-3}^{n-2} \widetilde{l}_1 \left[C_1(u,v)\right]^{t_1} \left[\left(u - C_1(u,v)\right)\left(v - C_1(u,v)\right)\right]^{n-2-t_1}\right. \\
&\times \left[\widehat{C}_1(1-u,1-v)\right]^{3-n+t_1} c_1(u,v) \\
&+ \sum_{t_4=0}^{1}\sum_{t_2=0}^{1}\sum_{t_1=\max(0,n-2-t_2-t_4)}^{\min(n-2-t_2,n-2-t_4)} \widetilde{l}_2 \left[C_1(u,v)\right]^{t_1} \left[u - C_1(u,v)\right]^{n-2-t_1-t_2} \\
&\times \left[v - C_1(u,v)\right]^{n-2-t_1-t_4} \left[\widehat{C}_1(1-u,1-v)\right]^{2-n+t_1+t_2+t_4} \\
&\times \left.\left[C_1^{\cdot,1}(u,v)\right]^{t_2} \left[1 - C_1^{\cdot,1}(u,v)\right]^{1-t_2} \left[C_1^{1,\cdot}(u,v)\right]^{t_4} \left[1 - C_1^{1,\cdot}(u,v)\right]^{1-t_4}\right\} dudv
\end{aligned} \tag{A.4}$$

Then, under $H_0 : C(u,v) = C_1(u,v) = uv$,

$$P\{A \mid H_0\} + P\{B \mid H_0\} \geq o(h_1) > 0,$$

where $o(h_1) = \alpha_n = \frac{n(n-1)^2(n+1)(n-2)!}{4^n A(n)} + \frac{\left(n^2-1\right)(n-1)!(n+1)!}{4^n (A(n))^2}$. It is clear that $P\{ND \mid H_0\} < 1$. Thus Theorem 2.2 is proved. □

Appendix B

Proof of Theorem 2.3. By using the similar probability integral transformation used in Equation (A.2), Equation (2.4) can be written as follows:

$$\begin{aligned}
\beta_n &= P\left\{H_0 \mid H_1\right\} \\
&= P\{\text{at least } n-1 \text{ of the sample values in } Z \text{ are in} \\
&\qquad \left(-\infty, X_{1:n}^{(2)}\right) \times \left(-\infty, Y_{1:n}^{(2)}\right) \mid H(x,y) = G(x,y)\Big\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\overline{G}(t,s)\right]^n g_{X_{n-1:n},Y_{n-1:n}}(t,s)\, dtds. \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\overline{G}(t,s)\right]^n \left(\sum_{t_1=n-3}^{n-2} \widetilde{l}_1 \left[G(t,s)\right]^{t_1} \left[\left(F_X(t) - G(t,s)\right)\left(F_Y(s) - G(t,s)\right)\right]^{n-2-t_1}\right. \\
&\quad \times \left[\overline{G}(t,s)\right]^{3-n+t_1} g(t,s) + \sum_{t_4=0}^{1}\sum_{t_2=0}^{1}\sum_{t_1=\max(0,n-2-t_2-t_4)}^{\min(n-2-t_2,n-2-t_4)} \widetilde{l}_2 \left[G(t,s)\right]^{t_1} \\
&\quad \times \left[\left(F_X(t) - G(t,s)\right)\right]^{n-2-t_1-t_2} \left[\left(F_Y(s) - G(t,s)\right)\right]^{n-2-t_1-t_4} \\
&\quad \times \left[\overline{G}(t,s)\right]^{2-n+t_1+t_2+t_4} \left[G^{\cdot,1}(t,s)\right]^{t_2} \left[f_Y(s) - G^{\cdot,1}(t,s)\right]^{1-t_2} \\
&\quad \times \left.\left[G^{1,\cdot}(t,s)\right]^{t_4} \left[f_X(t) - G^{1,\cdot}(t,s)\right]^{1-t_4}\right) dtds,
\end{aligned} \tag{A.5}$$

where $\widetilde{l}_1$ and $\widetilde{l}_2$ are

$$\begin{aligned}
\widetilde{l}_1 &= \frac{n!}{t_1!\left[(n-2-t_1)!\right]^2 (3-n+t_1)!} \\
\widetilde{l}_2 &= \frac{n!}{t_1!\,(n-2-t_1-t_2)!\,(n-2-t_1-t_4)!\,(2-n+t_1+t_2+t_4)!}.
\end{aligned}$$

By probability integral transformation we have the following equation.

$$\begin{aligned}
\beta_n &= \int_0^1\int_0^1 \left(\widehat{C}_2(1-u,1-v)\right)^n \left\{\sum_{t_1=n-3}^{n-2} \widetilde{l}_1 \left[C_2(u,v)\right]^{t_1} \left[\left(u - C_2(u,v)\right)\left(v - C_2(u,v)\right)\right]^{n-2-t_1}\right. \\
&\quad \times \left[\widehat{C}_2(1-u,1-v)\right]^{3-n+t_1} c_2(u,v) + \sum_{t_4=0}^{1}\sum_{t_2=0}^{1} \\
&\quad + \sum_{t_1=\max(0,n-2-t_2-t_4)}^{\min(n-2-t_2,n-2-t_4)} \widetilde{l}_2 \left[C_2(u,v)\right]^{t_1} \left[u - C_2(u,v)\right]^{n-2-t_1-t_2} \\
&\quad \times \left[v - C_2(u,v)\right]^{n-2-t_1-t_4} \left[\widehat{C}_2(1-u,1-v)\right]^{2-n+t_1+t_2+t_4} \\
&\quad \times \left.\left[C_2^{\cdot,1}(u,v)\right]^{t_2} \left[1 - C_2^{\cdot,1}(u,v)\right]^{1-t_2} \left[C_2^{1,\cdot}(u,v)\right]^{t_4} \left[1 - C_2^{1,\cdot}(u,v)\right]^{1-t_4}\right\} dudv.
\end{aligned}$$

Then, the probability of type II error under $H_1 : C(u,v) = C_2(u,v) = uv$ is

$$\beta_n = \int_0^1 \int_0^1 [(1-u)(1-v)]^n \Bigg((uv)^{n-2}(1-u)(1-v)\, n(n-1)^2 + (uv)^{n-2}(1-u)(1-v) \sum_{t_4=0}^{1} \sum_{t_2=0}^{1} \sum_{t_1=\max(0,n-2-t_2-t_4)}^{\min(n-2-t_2,n-2-t_4)} \widetilde{l}_2 \Bigg) dudv$$
$$= \frac{n(n-1)^2[(n-2)!]^2[(n+1)!]^2}{[(2n)!]^2} + \frac{(n-1)(n-1)!n![(n+1)!]^2}{[(2n)!]^2}.$$

Denote $K_n^{'} = \frac{n(n-1)^2[(n-2)!]^2[(n+1)!]^2}{[(2n)!]^2}$ and $K_n^{''} = \frac{(n-1)(n-1)!n![(n+1)!]^2}{[(2n)!]^2}$. Then,

$$\lim_{n\to\infty} K_n^{'} = \lim_{n\to\infty} \frac{n(n-1)^2[(n-2)!]^2[(n+1)!]^2}{[(2n)!]^2}$$
$$= \lim_{n\to\infty} \frac{n(n+1)^2[(n-1)!]^2(n!)^2}{[(2n)!]^2}$$
$$= \lim_{n\to\infty} \frac{n(n+1)^2[(n-1)!]^2}{4^n(A(n))^2}$$
$$= \lim_{n\to\infty} \frac{n(n+1)^2}{4^n} \cdot \lim_{n\to\infty} \left[\frac{(n-1)!}{A(n)}\right]^2$$
$$\lim_{n\to\infty} K_n^{'} = 0.$$

Similarly,

$$\lim_{n\to\infty} K_n^{''} = \lim_{n\to\infty} \frac{(n-1)(n-1)!n![(n+1)!]^2}{[(2n)!]^2}$$
$$= \lim_{n\to\infty} \frac{(n-1)(n+1)^2(n-1)!n!(n!)^2}{[(2n)!]^2}$$
$$= \lim_{n\to\infty} \frac{(n-1)(n+1)^2(n-1)!n!(n!)^2}{[(2n)(2n-1)(2n-2)\cdots 3.2.1]^2}$$
$$= \lim_{n\to\infty} \frac{(n-1)(n+1)^2(n-1)!n!}{4^n[(2n-1)(2n-3)(2n-5)\cdots 5.3.1]^2}$$
$$= \lim_{n\to\infty} \frac{(n-1)(n+1)^2}{4^n} \cdot \lim_{n\to\infty} \frac{(n-1)!n!}{[(2n-1)(2n-3)(2n-5)\cdots 5.3.1]^2}$$
$$= 0 \cdot \lim_{n\to\infty} \frac{(n-1)!n!}{[(2n-1)(2n-3)(2n-5)\cdots 5.3.1]^2}$$
$$= 0.$$

Thus $\lim_{n\to\infty}\beta_n = 0$.
Now consider $P\{ND \mid H_1\}$ under $H_1 : C(u,v) = C_2(u,v) = uv$,

$$P\{ND \mid H_1\} = 1 - P\{A \cup B \mid H_1\}$$
$$= 1 - P\{A \mid H_1\} - P\{B \mid H_1\}. \tag{A.6}$$

It is obvious that $\beta_n = P\{A \mid H_1\}$, then,

$$P\{A \mid H_1\} + P\{B \mid H_1\} \geq o(h_2) > 0,$$

where $o(h_2) = \frac{n(n-1)^2[(n-2)!]^2[(n+1)!]^2}{[(2n)!]^2} + \frac{(n-1)(n-1)!n![(n+1)!]^2}{[(2n)!]^2}$ and $P\{ND \mid H_1\} < 1$.Thus Theorem 2.3 is proved. □