Stat Methods Appl (2010) 19:541–565
DOI 10.1007/s10260-010-0147-7

A nonparametric symmetry test for absolutely continuous bivariate copulas

Arturo Erdely · José M. González-Barrios

Accepted: 6 September 2010 / Published online: 29 September 2010

Abstract Based on the works by Klement and Mesiar (Comment Math Univ Carolinae 47:141–148, 2006) and Nelsen (Stat Pap 48:329–336, 2007) on maximal asymmetry of copulas, we define and study the concept of tri-symmetry and we propose a simple statistic to test symmetry of a bivariate copula, given a random sample of an absolutely continuous bivariate random vector. We also make a power comparison against some other well known nonparametric symmetry tests.

Keywords Symmetry · Nonparametric exact test · Asymmetric copulas

1 Introduction

In recent papers the problem of finding extreme nonexchangeable copulas or extreme asymmetric copulas has been studied, see Klement and Mesiar (2006), Nelsen (2007). In these papers the following result is proved:

A. Erdely (✉)
Programa de Actuaría, Facultad de Estudios Superiores Acatlán, Universidad Nacional Autónoma de México, Av. Alcanfores y San Juan Totoltepec s/n, Santa Cruz Acatlán, Naucalpan, 53150 Estado de México, México
e-mail: arturoerdely@yahoo.com.mx

J. M. González-Barrios
Departamento de Probabilidad y Estadística, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México,
Circuito Escolar s/n, Ciudad Universitaria, 04510 México, DF, México
e-mail: gonzaba@sigma.iimas.unam.mx

Lemma 1 *Let C be a copula on I^2, where $I = [0, 1]$, then*

$$\sup_{(u,v)\in I^2} |C(u, v) - C(v, u)| \leq \frac{1}{3},$$

and the inequality is best possible.

In fact, in Klement and Mesiar (2006), Nelsen (2007) we can find examples in which the upper bound $\frac{1}{3}$ is attained. The copulas for which this upper bound is attained will be called *maximally nonexchangeable or extremely asymmetric*. Even more, in Nelsen (2007), Theorem 3.1, we can find an interesting characterization of the familiy $\mathcal{C}$ of maximally nonexchangeable or extremely asymmetric copulas, that is,

Theorem 1 *If we define the following sets of copulas:*

$$\mathcal{C}_1 = \{C\,|C(2/3, 1/3) = 0\}, \quad \mathcal{C}_2 = \{C\,|C(1/3, 2/3) = 1/3\},$$
$$\mathcal{C}_3 = \{C\,|C(1/3, 2/3) = 0\}, \quad \mathcal{C}_4 = \{C\,|C(2/3, 1/3) = 1/3\},$$

then $\mathcal{C} = (\mathcal{C}_1 \cap \mathcal{C}_2) \cup (\mathcal{C}_3 \cap \mathcal{C}_4)$, i.e., $C \in \mathcal{C}$ if and only if either $C(1/3, 2/3) = 1/3$ and $C(2/3, 1/3) = 0$, or $C(1/3, 2/3) = 0$ and $C(2/3, 1/3) = 1/3$.

In Durante et al. (2008a), several measures of non-exchangeability are analyzed. If we let C to be a copula, we are interested in the following measure of nonexchangeability or asymmetry:

$$\delta(C) = 3 \sup_{(u,v)\in I^2} |C(u, v) - C(v, u)|. \tag{1}$$

Then, we have the following properties of $\delta(C)$:

(i) $\delta(C) = 0$ if and only if C is exchangeable or symmetric.
(ii) For any copula C, $0 \leq \delta(C) \leq 1$.
(iii) C is maximally nonexchangeable or extremely asymmetric if and only if $\delta(C) = 1$.
(iv) From Theorem 1, the value of $\delta(C) = 1$ for maximally nonexchangeable or extremely asymmetric copulas is attained at $(u, v) = (1/3, 2/3)$ or at $(u, v) = (2/3, 1/3)$.

In Sect. 2 we construct a family of absolutely continuous copulas which includes the product copula and a maximally nonexchangeable copula. We also study some properties of copulas C such that $C(2/3, 1/3) = C(1/3, 2/3)$ and define the concept of *tri-symmetry* of copulas.

In Sect. 3 we propose a simple statistic to test for tri-symmetry of copulas, studying its exact distribution as well as its asymptotic distribution.

In Sect. 4 we include several simulations in order to check the power of the proposed test against some other well known nonparametric symmetry tests.

2 Construction of a family of copulas and definition of tri-symmetry

In this section we first construct a family of absolutely continuous copulas which includes the product copula Π, and also includes a maximal nonexchangeable copula.

Let $0 \le \epsilon \le 6$ and define

$$c_\epsilon(u, v) = \begin{cases} \epsilon/6 & \text{if} \quad (u, v) \in I_1 \cup I_2 \cup I_5 \cup I_6 \cup I_7 \cup I_9 \\ 3 - \epsilon/3 & \text{if} \quad (u, v) \in I_3 \cup I_4 \cup I_8, \end{cases} \tag{2}$$

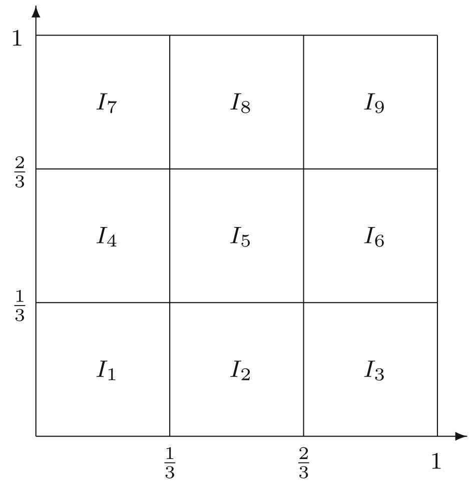
where $I_1 = (0, 1/3] \times (0, 1/3]$, $I_2 = (1/3, 2/3] \times (0, 1/3]$, $I_3 = (2/3, 1] \times (0, 1/3]$, $I_4 = (0, 1/3] \times (1/3, 2/3]$, $I_5 = (1/3, 2/3] \times (1/3, 2/3]$, $I_6 = (2/3, 1] \times (1/3, 2/3]$, $I_7 = (0, 1/3] \times (2/3, 1]$, $I_8 = (1/3, 2/3] \times (2/3, 1]$ and $I_9 = (2/3, 1] \times (2/3, 1]$, see Fig. 1.

Then $c_\epsilon(u, v)$ is a nonnegative Borel measurable function, which is bounded. Define for $(u, v) \in [0, 1]^2$, $C_\epsilon(u, v) = \int_0^v \int_0^u c_\epsilon(s, t) ds dt$, then integrating and simplifying

$$C_\epsilon(u, v) =: \\ \times \begin{cases} uv\epsilon/6 & \text{in } I_1 \cup I_2 \\ v\epsilon/9 + (u - 2/3)v(3 - \epsilon/3) & \text{in } I_3 \\ u\epsilon/18 + u(v - 1/3)(3 - \epsilon/3) & \text{in } I_4 \\ u\epsilon/18 + (v - 1/3)(1 - \epsilon/9 + (u - 1/3)\epsilon/6) & \text{in } I_5 \\ \epsilon/27 + (u - 2/3)(1 - \epsilon/9) + (v - 1/3)\{1 - \epsilon/9 + (u - 1/3)\epsilon/6\} & \text{in } I_6 \\ u\{1 - \epsilon/18 + (v - 2/3)\epsilon/6\} & \text{in } I_7 \\ u\epsilon/9 + 1/3 - \epsilon/18 + (v - 2/3)\{\epsilon/18 + (u - 1/3)(3 - \epsilon/3)\} & \text{in } I_8 \\ u(1 - \epsilon/18) + \epsilon/18 - 1/3 + (v - 2/3)\{1 - \epsilon/18 + (u - 2/3)\epsilon/6\} & \text{in } I_9. \end{cases}$$

From this expression, it is easy to check that C_ϵ satisfies the boundary conditions of a copula, and that $C_\epsilon(1, 1) = 1$. Since $c_\epsilon(s, t) \ge 0$ for every $(s, t) \in [0, 1]^2$, then C_ϵ is obviously two-increasing. Therefore, for any $0 \le \epsilon \le 6$, C_ϵ is an absolutely

Fig. 1 Regions $I_1, I_2, \ldots, I_9$ for copula density (2)



continuous copula. In fact, if $\epsilon = 6$ it is clear that $C_6(u, v) = uv$ for every $u, v \in [0, 1]$, since $c_6(s, t) = \frac{\partial}{\partial t}\frac{\partial}{\partial s}st = \frac{\partial}{\partial t}\frac{\partial}{\partial s}C_\pi(s, t) = 1$ for every $(s, t) \in [0, 1]^2$, that is, C_6 is the product copula Π.

Hence, the family of copulas C_ϵ with $0 \leq \epsilon \leq 6$, includes the product (or independence) copula, and for any $\epsilon \geq 0$ it is absolutely continuous; when $\epsilon = 0$, C_0 is maximally nonexchangeable or extremely asymmetric.

Of course if we define for $0 \leq \epsilon \leq 6$,

$$c'_\epsilon(u, v) = \begin{cases} \epsilon/6 & \text{if} \quad (u, v) \in I_1 \cup I_3 \cup I_4 \cup I_5 \cup I_8 \cup I_9 \\ 3 - \epsilon/3 & \text{if} \quad (u, v) \in I_2 \cup I_6 \cup I_7, \end{cases}$$

where $I_1, \ldots, I_9$ are defined as above, and we define $C'_\epsilon(u, v) = \int_0^v \int_0^u c'_\epsilon(s, t)dsdt$, then $C'_\epsilon(u, v)$ is the symmetric version of $C_\epsilon(u, v)$ defined above.

Besides, the family of copulas C'_ϵ with $0 \leq \epsilon \leq 6$, includes the independence copula for $\epsilon = 6$, and for any $\epsilon \geq 0$ it is absolutely continuous; when $\epsilon = 0$, C'_0 is maximally nonexchangeable or extremely asymmetric.

On the other hand it is easy to check that

$$C_\epsilon(2/3, 1/3) = \frac{\epsilon}{27} = C'_\epsilon(1/3, 2/3)$$

and

$$C_\epsilon(1/3, 2/3) = \frac{1}{3} - \frac{\epsilon}{54} = C'_\epsilon(2/3, 1/3).$$

Therefore, if $\epsilon = 0$ we have that C_0 belongs to $(\mathcal{C}_1 \cap \mathcal{C}_2)$ and C'_0 belongs to $(\mathcal{C}_3 \cap \mathcal{C}_4)$, as defined in Theorem 1.

Besides, it is easy to see that the measure on nonexchangeability defined in (1) is such that

$$\delta(C_\epsilon) = 3|C_\epsilon(2/3, 1/3) - C_\epsilon(1/3, 2/3)| = 1 - \frac{\epsilon}{6}$$

and

$$\delta(C'_\epsilon) = 3|C'_\epsilon(2/3, 1/3) - C'_\epsilon(1/3, 2/3)| = 1 - \frac{\epsilon}{6}.$$

Let us observe that in Theorem 1, given a copula C, the value of $|C(2/3, 1/3) - C(1/3, 2/3)|$ determines if C is maximally nonexchangeable or extremely asymmetric. We will prove a general result involving this quantity.

Proposition 1 *Let $C(u, v)$ be the copula associated to the random vector (U, V) with uniform $(0, 1)$ marginals, such that*

$$|C(2/3, 1/3) - C(1/3, 2/3)| = 0. \tag{3}$$

Define I_j, $j = 1, 2, \ldots, 9$, *as in* (2), *and let* $p_j = P(\{(U, V) \in I_j\})$, *for* $j = 1, \ldots, 9$. *Then*

$$p_2 = p_4, \quad p_3 = p_7 \quad \textit{and} \quad p_6 = p_8.$$

Proof Let C be a copula such that (3) holds. Since C is a continuous joint distribution function of a random vector (U, V) with $U(0, 1)$ marginals, then for every $(u, v) \in I^2$, $C(u, v) = P(\{(U, V) \in (0, u] \times (0, v]\})$. Since $(0, 2/3] \times (0, 1/3] = I_1 \cup I_2$ and $(0, 1/3] \times (0, 2/3] = I_1 \cup I_4$ (see Fig. 1), and the sets I_j are pairwise disjoint, then

$$0 = |C(2/3, 1/3) - C(1/3, 2/3)| = |P(I_1 \cup I_2) - P(I_1 \cup I_4)| = |p_2 - p_4|,$$

and $p_2 = p_4$. Since $C(1, 1/3) = C(1/3, 1) = 1/3$, then $1/3 = p_1 + p_2 + p_3 = p_1 + p_4 + p_7$, so $p_3 = p_7$. Finally, observe that

$$\begin{aligned} p_8 + p_9 &= P(\{(U, V) \in (1/3, 1] \times (2/3, 1]\}) \\ &= C(1, 1) - C(1/3, 1) - C(2/3, 1) + C(1/3, 2/3) \\ &= 1 - 1/3 - 2/3 + C(1/3, 2/3) \end{aligned}$$

and

$$\begin{aligned} p_6 + p_9 &= P(\{(U, V) \in (2/3, 1] \times (1/3, 1]\}) \\ &= C(1, 1) - C(2/3, 1) - C(1/3, 1) + C(2/3, 1/3) \\ &= 1 - 2/3 - 1/3 + C(2/3, 1/3). \end{aligned}$$

Therefore

$$|(p_8 + p_9) - (p_6 + p_9)| = |C(1/3, 2/3) - C(2/3, 1/3)| = 0.$$

Hence $p_6 = p_8$. □

A copula C satisfying condition (3) of the last Proposition will be called *tri-symmetric*. Here we observe that any *symmetric copula* C, that is, $C(u, v) = C(v, u)$ for every $u, v \in I$, is obviously tri-symmetric. Hence, every *Archimedean copula* is also tri-symmetric, see definition of Archimedean copula in Chapter 4, and also Theorem 4.1.5 in Nelsen (2006).

Now, if we know that C is a tri-symmetric copula, how large can be the value of $\delta(C)$ defined in Eq. (1)? To solve this question we have the following:

Proposition 2 *Let* C *be a tri-symmetric copula. Then*

$$\delta(C) = 3 \sup_{(u,v)\in I^2} |C(u, v) - C(v, u)| \leq \frac{1}{2}, \tag{4}$$

and the inequality is best possible.

Proof If C is tri-symmetric then $C(1/3, 2/3) = C(2/3, 1/3)$. Let $q = C(1/3, 2/3) = C(2/3, 1/3)$, then $0 \leq q \leq 1/3$ using Frechet-Hoeffding lower and upper bounds, see Nelsen (2006). Recall the regions defined in the construction of the family of copulas given in (2), see Fig. 1.

First assume that $q = 0$, by Proposition 1 we know that $p_3 = p_7$, but since $q = p_1+p_2 = p_1+p_4$ and $p_1+p_2+p_3 = p_1+p_4+p_7 = 1/3$, then $p_3 = p_7 = 1/3$. This also implies that $p_6 = p_9 = p_8 = 0$ and $p_5 = 1/3$, that is, the regions with masses $1/3$ are I_3, I_5 and I_7, see Fig. 1. Hence C is an ordinal sum of copulas, but on the secondary diagonal instead of the main diagonal, see Schweizer and Sklar (2005). C can also be described as a shuffle of copulas, see Durante et al. (2009). Now, let C_1 and C_2 be two copulas we want to find an upper bound for

$$\delta(C_1, C_2) = \sup_{(u,v)\in[0,1]^2} |C_1(u, v) - C_2(v, u)|.$$

Using the Fréchet-Hoeffding bounds, we know that $W(u, v) \leq C_i(u, v) \leq M(u, v)$, for $i = 1, 2$ where $W(u, v) = \max\{u + v - 1, 0\}$ and $M(u, v) = \min\{u, v\}$ for every $(u, v) \in [0, 1]^2$. Then

$$\begin{aligned}
\delta(C_1, C_2) &= \sup_{(u,v)\in[0,1]^2} |C_1(u, v) - C_2(v, u)| \\
&\leq \sup_{(u,v)\in[0,1]^2} |M(u, v) - W(v, u)| \\
&= \sup_{(u,v)\in[0,1]^2} (M(u, v) - W(u, v)) \\
&\leq M(1/2, 1/2) - W(1/2, 1/2) \\
&= \frac{1}{2},
\end{aligned}$$

where we used the symmetry of W, and the definitions of W and M to obtain the supremum of $M - W$. Therefore, for every $(u, v) \in I_3$, since $(v, u) \in I_7$, we have that

$$\sup_{(u,v)\in I_3} |C(u, v) - C(v, u)| \leq \frac{1}{3} \sup_{(u,v)\in[0,1]^2} |C_1(u, v) - C_2(v, u)| \leq \frac{1}{6},$$

where C_1 is the copula associated to the ordinal sum or shuffle in I_3 and C_2 is the copula associated in I_7.

We know by Theorem 1 from the introduction, that for any copula C_3, $\sup_{(u,v)\in[0,1]^2} |C_3(u, v) - C_3(v, u)| \leq 1/3$. So, if we let C_3 be the copula associated to C in the ordinal sum or shuffle associated to region I_5, we have that

$$\sup_{(u,v)\in I_5} |C(u, v) - C(v, u)| \leq \frac{1}{3} \sup_{(u,v)\in[0,1]^2} |C_3(u, v) - C_3(v, u)| \leq \frac{1}{9}.$$

Therefore, in the case that $q = 0$, we obtain

$$\delta(C) = 3 \sup_{(u,v)\in[0,1]^2} |C(u, v) - C(v, u)| \leq 3 \cdot \frac{1}{6} = \frac{1}{2}.$$

Lemma 2.1 in Nelsen (2007) states that for any copula C and $u, v \in [0, 1]$,

$$|C(u, v) - C(v, u)| \leq \min\{u, v, 1 - u, 1 - v, |u - v|\}.$$

Let C be a copula, we want to find upper bounds for

$$\sup_{(u,v)\in I_1} |C(u, v) - C(v, u)| \quad \text{and} \quad \sup_{(u,v)\in I_9} |C(u, v) - C(v, u)|.$$

Using Lemma 2.1 in Nelsen (2007) we have that

$$\sup_{(u,v)\in I_1} \leq \min\{u, v, 1 - u, 1 - v, |v - u|\} = \min\{u, v, |u - v|\} = \frac{1}{9}, \tag{5}$$

which is attained at $u = 1/9$ and $v = 2/9$, or $u = 2/9$ and $v = 1/9$. By a totally analogous argument, we obtain that

$$\sup_{(u,v)\in I_9} |C(u, v) - C(v, u)| \leq \frac{1}{9}. \tag{6}$$

Second, assume that $q = 1/3$, with $p_2 = p_4 = 1/3$, this implies that $p_1 = p_3 = p_5 = p_6 = p_7 = p_8 = 0$ and $p_9 = 1/3$. In this case, C is the shuffle, see Durante et al. (2009), of three copulas in the regions I_2, I_4 and I_9. Using exactly the same arguments as in the case $q = 0$, we have that

$$\sup_{(u,v)\in I_2} |C(u, v) - C(v, u)| \leq \frac{1}{6}.$$

Together with Eq. (6), we obtain again

$$\delta(C) = 3 \sup_{(u,v)\in[0,1]^2} |C(u, v) - C(v, u)| \leq 3 \cdot \frac{1}{6} = \frac{1}{2}.$$

The case $q = 1/3$ with $p_1 = 1/3$ and $p_6 = p_8 = 1/3$ can be analized exactly in the same way.

If $q = 1/3$ with $p_1 = p_5 = p_9 = 1/3$ the result follows from (5) and (6) and observing that in this case C is simply the usual ordinal sum of three copulas in the main diagonal. Note that in this case $\delta(C) \leq 1/9$.

We have analized all extreme cases, for the remaining cases some of the p_i 's are greater than zero but less than $1/3$. It can be seen that in these cases the upper bound still applies. □

To see that the inequality is best possible, we give six examples in which the upper bound in (4) is attained. Let $C_i, i = 1, 2, \ldots, 6$ be shuffles of M, for definitions see Nelsen (2006). The support of C_1 consists of two line segments in $[0, 1]^2$, one connecting $(0, 2/3)$ to $(1/3, 1)$ and one connecting $(1/3, 2/3)$ to $(1, 0)$; the support of C_2 consists of two lines, one connecting $(0, 1)$ to $(2/3, 1/3)$ and another one connecting $(2/3, 0)$ to $(1, 1/3)$, see Fig. 2. Observe that in these cases $C_1(1/3, 2/3) = C_1(2/3, 1/3) = C_2(1/3, 2/3) = C_2(2/3, 1/3) = 0$, so C_1 and C_2 are tri-symmetric. Besides

$$|C_i(1/6, 5/6) - C_i(5/6, 1/6)| = \frac{1}{6} \quad \text{for} \quad i = 1, 2.$$

Hence $\delta(C_1) = \delta(C_2) = 1/2$.

The support of C_3 consists of three line segments in $[0, 1]^2$, one connecting $(0, 1/3)$ to $(1/3, 2/3)$, another connecting $(1/3, 1/3)$ to $(2/3, 0)$, and a third one connecting $(2/3, 2/3)$ to $(1, 1)$; the support of C_4 consists of three line segments one connecting $(0, 2/3)$ to $(1/3, 1/3)$, another connecting $(1/3, 0)$ to $(2/3, 1/3)$ and a third one connecting $(2/3, 2/3)$ to $(1, 1)$, see Fig. 3. Observe that in these cases $C_3(1/3, 2/3) = C_3(2/3, 1/3) = C_4(1/3, 2/3) = C_4(2/3, 1/3) = 1/3$, so C_3 and C_4 are tri-symmetric. Besides

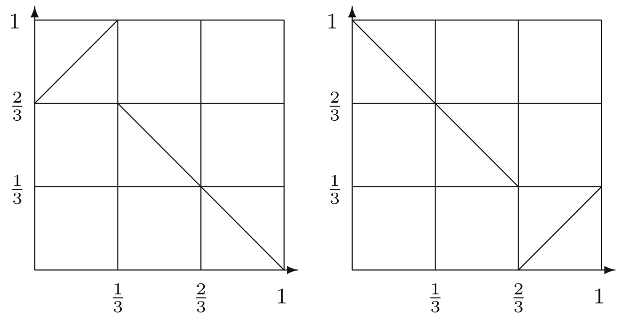


Fig. 2 The supports of C_1 and C_2

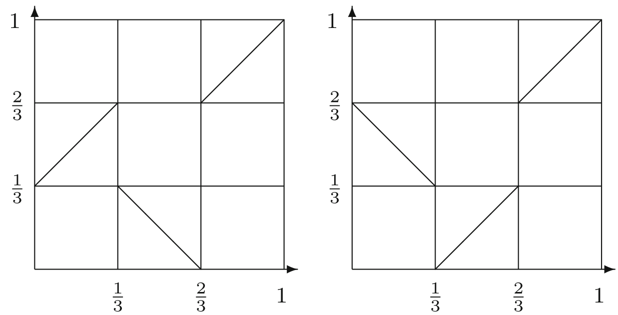


Fig. 3 The supports of C_3 and C_4

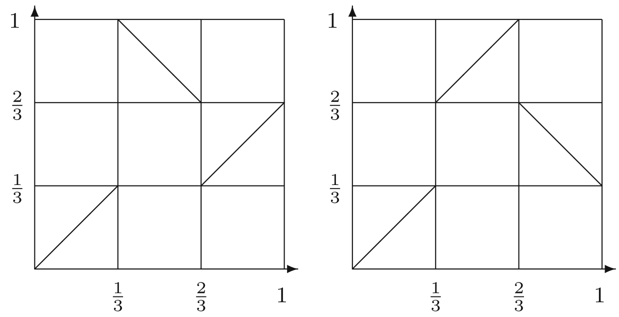


Fig. 4 The supports of C_5 and C_6

$$|C_i(1/2, 1/6) - C_i(1/6, 1/2)| = \frac{1}{6} \quad \text{for} \quad i = 3, 4.$$

Hence $\delta(C_3) = \delta(C_4) = 1/2$.

The support of C_5 consists of three line segments in $[0, 1]^2$, one connecting $(0, 0)$ to $(1/3, 1/3)$, another connecting $(1/3, 1)$ to $(2/3, 2/3)$, and a third one connecting $(2/3, 1/3)$ to $(1, 2/3)$; the support of C_6 consists of three line segments one connecting $(0, 0)$ to $(1/3, 1/3)$, another connecting $(1/3, 2/3)$ to $(2/3, 1)$, and a third one connecting $(2/3, 2/3)$ to $(1, 1/3)$, see Fig. 4. Observe that in these cases $C_5(1/3, 2/3) = C_5(2/3, 1/3) = C_6(1/3, 2/3) = C_6(2/3, 1/3) = 1/3$, so C_5 and C_6 are tri-symmetric. Besides

$$|C_i(1/2, 5/6) - C_i(5/6, 1/2)| = \frac{1}{6} \quad \text{for} \quad i = 5, 6.$$

Hence $\delta(C_5) = \delta(C_6) = 1/2$.

3 A test for tri-symmetry of copulas

In this section we will propose a simple statistic that can be used to test the following hypotheses for any absolutely continuous copula C :

$$H_0 : \ C \text{ is a tri-symmetric copula} \quad \text{vs} \quad H_1 : \ C \text{ is not a tri-symmetric copula} \tag{7}$$

Let us assume that we have a random sample $\mathbf{X} = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ of size n of a continuous random vector (X, Y) in $\mathbb{R}^2$, with underlying copula $C_{X,Y}$. We will assume without loosing generality that $X_1 < X_2 < \cdots < X_n$. Recall that the empirical copula, see Nelsen (2006), is defined by

$$C_n(i/n, j/n) = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{\text{rank}(X_k) \leq i, \quad \text{rank}(Y_k) \leq j\}}$$

$$= \frac{1}{n}\sum_{k=1}^{i}\mathbf{1}_{\{\mathrm{rank}(Y_k)\le j\}},$$

where $i, j \in \{1, 2, \ldots, n\}$ and the rank is defined as usual. We also define $C_n(0, j/n) = C_n(i/n, 0) = 0$ for every $i, j \in \{1, 2, \ldots, n\}$. Then C_n is a subcopula, see Nelsen (2006), with domain $\{0, 1/n, \ldots, (n-1)/n, 1\}\times\{0, 1/n, \ldots, (n-1)/n, 1\}$. It is well known, see Deheuvels (1979), that

$$\lim_{n\to\infty}\sup_{(i,j)\in\{1,2,\ldots,n\}}|C_n(i/n, j/n) - C_{X,Y}(i/n, j/n)| = 0 \quad \text{almost surely } [P],$$

where P is the probability measure defined on $\mathbb{R}^2$ by the bivariate distribution function $C_{X,Y}$. Therefore, if n is large enough C_n is a very good approximation of the real $C_{X,Y}$. We will use this important asymptotic property, to define an adequate statistic to test (7). But first, in order to make sense for each $n \ge 1$ of our next definition, we extend the domain of the empirical copula C_n to be $[0, 1]^2$ in the following way

$$C_n(u,v) = \begin{cases} C_n\left(\frac{i}{n}, \frac{j}{n}\right) & \text{if } (u,v) \in \left[\frac{i}{n}, \frac{(i+(i+1))}{2n}\right) \times \left[\frac{j}{n}, \frac{(j+(j+1))}{2n}\right) \\ C_n\left(\frac{i}{n}, \frac{j+1}{n}\right) & \text{if } (u,v) \in \left[\frac{i}{n}, \frac{(i+(i+1))}{2n}\right) \times \left[\frac{(j+(j+1))}{2n}, \frac{(j+1)}{n}\right) \\ C_n\left(\frac{i+1}{n}, \frac{j}{n}\right) & \text{if } (u,v) \in \left[\frac{(i+(i+1))}{2n}, \frac{(i+1)}{n}\right) \times \left[\frac{j}{n}, \frac{(j+(j+1))}{2n}\right) \\ C_n\left(\frac{i+1}{n}, \frac{j+1}{n}\right) & \text{if } (u,v) \in \left[\frac{(i+(i+1))}{2n}, \frac{(i+1)}{n}\right) \times \left[\frac{(j+(j+1))}{2n}, \frac{(j+1)}{n}\right), \end{cases}$$

for some $i, j \in \{0, 1, \ldots, n-1\}$.

Definition 1 Let $\mathbf{X} = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ be a random sample of size n of a continuous random vector (X, Y) in $\mathbb{R}^2$, with copula $C_{X,Y}$. We define $S_n = |T_n|$, where

$$T_n = C_n(1/3, 2/3) - C_n(2/3, 1/3). \tag{8}$$

Recall that a copula C is tri-symmetric if and only if $C(1/3, 2/3) = C(2/3, 1/3)$, then using the asymptotic property of the empirical copula, an appropriate statistic for tri-symmetry is S_n. In the rest of this section we will study the exact distribution of T_n and its asymptotic distribution under independence. To do so, we start with some general results.

Theorem 2 *Let $\mathbf{X} = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ be a random sample of size n of a continuous random vector (X, Y) of independent random variables in $\mathbb{R}^2$. Let i be a fixed integer such that $1 \le i < n/2$. Let us define the random variables*

$$Z_n = n\,C_n(i/n, i/n), \quad V_n = n\,C_n(i/n, 1 - i/n), \quad W_n = n\,C_n(1 - i/n, i/n), \tag{9}$$

where C_n is the empirical copula. Let

$$T_{i,n} = V_n - W_n \tag{10}$$

then

$$\mathbb{P}(T_{i,n} = t) = \sum_{v=0}^{i} \mathbb{P}(V_n = v, W_n = v - t), \quad t \in \{-i, -i+1, \ldots, -1, 0, 1, \ldots, i\}, \tag{11}$$

where

$$\mathbb{P}(V_n = v, W_n = w) = \sum_{z=0}^{i} \frac{\binom{i-z}{v-z}\binom{n+z-2i}{v}}{\binom{n-i}{i}} \frac{\binom{i-z}{w-z}\binom{n+z-2i}{w}}{\binom{n-i}{i}} \frac{\binom{i}{z}\binom{n-i}{i-z}}{\binom{n}{i}}, \tag{12}$$

for $z \in \{0, 1, \ldots, i\}$ *and* $v, w \in \{z, z+1, \ldots \min\{i, n-2i+z\}\}$.

Proof Let $1 \leq i < n/2$ be a fixed integer and let $\mathbf{X} = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ be a random sample of size n of a continuous random vector (X, Y) of independent random variables, without loosing generality we will assume that $X_1 < X_2 < \cdots < X_n$. Since $1 \leq i < n/2$, then $1 \leq i < n/2 < n-i \leq n-1$. From the definition of the empirical copula C_n, Z_n is the number of pairs (X_j, Y_j) such that $j \leq i$ and $\text{rank}(Y_j) \leq i$, V_n is the number of pairs (X_j, Y_j) such that $j \leq i$ and $\text{rank}(Y_j) \leq n-i$, and W_n is the number of pairs (X_j, Y_j) such that $j \leq n-i$ and $\text{rank}(Y_j) \leq i$. Then $Z_n \leq \min\{V_n, W_n\}$.

First we will obtain the joint density of (V_n, W_n, Z_n), by conditioning

$$\begin{aligned}\mathbb{P}(V_n = v, W_n = w, Z_n = z) &= \mathbb{P}(V_n = v, W_n = w \mid Z_n = z)\, \mathbb{P}(Z_n = z) \\ &= \mathbb{P}(V_n = v \mid Z_n = z)\, \mathbb{P}(W_n = w \mid Z_n = z)\, \mathbb{P}(Z_n = z)\end{aligned} \tag{13}$$

where the second equality follows from conditional independence of V_n and W_n given $\{Z_n = z\}$. The values of $\text{rank}(Y_1), \text{rank}(Y_2), \ldots, \text{rank}(Y_n)$ are simply a random permutation in the set $\{1, 2, \ldots, n\}$, and from the independence assumption, each of the $n!$ permutations has the same probability.

Now, we will show that Z_n has hypergeometric distribution

$$\mathbb{P}(Z_n = z) = \frac{\binom{i}{z}\binom{n-i}{i-z}}{\binom{n}{i}}, \quad z \in \{0, 1, \ldots, i\}.$$

In order to do so, we observe that $0 \leq Z_n \leq i$, and that $Z_n = z$ if and only if the condition $\text{card}\{\text{rank}(Y_j) \leq i \mid j = 1, 2, \ldots, i\} = z$ holds. Since $nC_n(i/n, n/n) = i$, using combinatorial arguments it is easy to see that

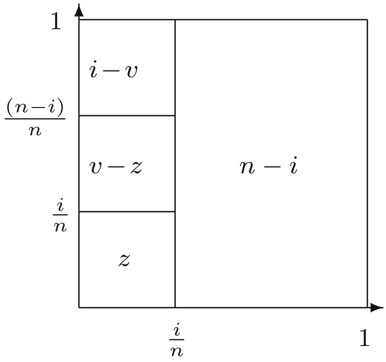


Fig. 5 Number of points by region induced by the empirical copula

$$
\begin{aligned}
\mathbb{P}(Z_n = z) &= \frac{\binom{i}{z}[i\cdots(i-z+1)][(n-i)\cdots(n-i-(i-z)+1)][(n-i)\cdots 1]}{n!} \\
&= \frac{\binom{i}{z}\binom{n-i}{i-z}}{\binom{n}{i}}. \qquad (14)
\end{aligned}
$$

In order to find the conditional density $h(v, z) = \mathbb{P}(V_n = v | Z_n = z)$, we observe that $z \le v \le i$. In Fig. 5 we give the number of points in each of the regions induced by the empirical copula, such that, $Z_n = z$ and $V_n = v$.

By using combinatorial arguments again, and the expression for $\mathbb{P}(Z_n = z)$, we obtain

$$
\begin{aligned}
h(v, z) &= \frac{\mathbb{P}(Z_n = z, V_n = v)}{\mathbb{P}(Z_n = z)} \\
&= \frac{\binom{i}{z}[i\cdots(i-z+1)]\binom{i-z}{v-z}[(n-2i)\cdots(n-2i-(v-z)+1)][i\cdots(i-(i-v)+1)](n-i)!}{n!\,\mathbb{P}(Z_n = z)} \\
&= \frac{\binom{i-z}{v-z}\binom{n+z-2i}{v}}{\binom{n-i}{i}}.
\end{aligned}
$$

Hence,

$$
\mathbb{P}(V_n = v \mid Z_n = z) = \frac{\binom{i-z}{v-z}\binom{n+z-2i}{v}}{\binom{n-i}{i}} \quad \text{if} \quad v \in \{z, z+1, \ldots, \min\{i, n+z-2i\}\}. \qquad (15)
$$

Using similar arguments we find that the conditional density of W_n given that $Z_n = z$ is given by

$$\mathbb{P}(W_n = w \mid Z_n = z) = \frac{\binom{i-z}{w-z}\binom{n+z-2i}{w}}{\binom{n-i}{i}}$$

$$\text{if} \quad w \in \{z, z+1, \ldots, \min\{i, n+z-2i\}\}. \tag{16}$$

Therefore, from Eqs. (13), (14), (15) and (16)

$$\mathbb{P}(V_n = v, W_n = w) = \sum_{z=0}^{i} \mathbb{P}(V_n = v, W_n = w, Z_n = z),$$

which is Eq. (12). Finally, for $T_{i,n} = V_n - W_n$ and summing over all indices we get

$$\mathbb{P}(T_{i,n} = t) = \sum_{v=0}^{i} \mathbb{P}(V_n = v, W_n = v - t), \quad t \in \{-i, -i+1, \ldots, -1, 0, 1, \ldots, i\},$$

which is Eq. (11). □

Corollary 1 *Let* $\mathbf{X} = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ *be a random sample of size* n *of a continuous random vector* (X, Y) *of independent random variables in* $\mathbb{R}^2$. *Let* i *be a fixed integer such that* $1 \leq i < n/2$. *Let us define the random variables* V_n, W_n *and* $T_{i,n}$ *as in Theorem 2. If we let* $S_{i,n} = |T_{i,n}|$, *then*

$$\mathbb{P}(S_{i,n} = s) = \begin{cases} \mathbb{P}(T_{i,n} = 0) & \text{if} \quad s = 0, \\ 2\mathbb{P}(T_{i,n} = s) & \text{if} \quad s = 1, 2, \ldots, i. \end{cases} \tag{17}$$

Proof It is enough to observe that by Eq. (12) the joint density of V_n and W_n is symmetrical, hence $T_{i,n}$ is a symmetric random variable, and the result follows from Eq. (11). □

Now we will obtain the expectation and variance of $T_{i,n}$ given in Eq. (10), using an alternative definition.

Proposition 3 *Let* $\mathbf{X} = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ *be a random sample of size* n *of a continuous random vector* (X, Y) *of independent random variables in* $\mathbb{R}^2$. *Let* i *be a fixed integer such that* $1 \leq i < n/2$. *Let us define the random variables* V_n, W_n *and* $T_{i,n}$ *as in Theorem 2. Then*

$$E(T_{i,n}) = 0 \quad \text{and} \quad \text{Var}(T_{i,n}) = \frac{2i^2(n-2i)}{n(n-1)}. \tag{18}$$

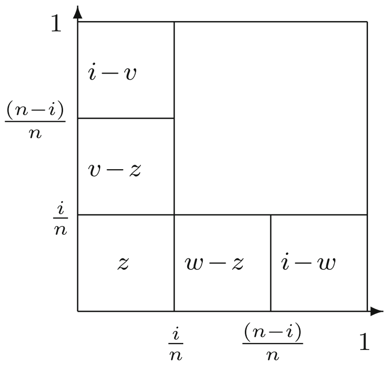


Fig. 6 The difference $T_{i,n} = V_n - W_n$ as in (10) or (19)

Proof Let $1 \leq i < n/2$ be a fixed integer and let $\mathbf{X} = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ be a random sample of size n of a continuous random vector (X, Y) of independent random variables, without loosing generality we will assume that $X_1 < X_2 < \cdots < X_n$. We know that $T_{i,n} = V_n - W_n = n\, C_n(i/n, 1 - i/n) - n\, C_n(1 - i/n, i/n)$, then $T_{i,n}$ is the difference of the number of pairs (X_j, Y_j) such that $j \leq i$ and $\text{rank}(Y_j) \leq n - i$ and the number of pairs (X_j, Y_j) such that $j \leq n - i$ and $\text{rank}(Y_j) \leq i$, see Fig. 6. Equivalently,

$$T_{i,n} = \sum_{j=1}^{i} \mathbf{1}_{A_j} - \sum_{j=i+1}^{n-i} \mathbf{1}_{B_j}, \tag{19}$$

where A_j and B_j are the events

$$A_j = \{\text{rank}(Y_j) \in \{i+1, i+2, \ldots, n - i\}\} \quad \text{and} \quad B_j = \{\text{rank}(Y_j) \in \{1, 2, \ldots, i\}\}.$$

Since, under independence, the ranks of the $Y_j's$ are permutations equally probable in the set $\{1, 2, \ldots, n\}$. Then $E(\mathbf{1}_{A_j}) = P(A_j) = (n-2i)/n$ for every $j \in \{1, 2, \ldots, i\}$, and $E(\mathbf{1}_{B_j}) = P(B_j) = i/n$ for every $j \in \{i + 1, i + 2, \ldots, n - i\}$. Besides,

$$E(\mathbf{1}_{A_j}\mathbf{1}_{A_k}) = \begin{cases} \frac{(n-2i)}{n}\frac{(n-2i-1)}{(n-1)} & \text{if} \quad j \neq k,\ j,\ k \in \{1, \ldots, i\} \\ \frac{n-2i}{n} & \text{if} \quad j = k,\ j \in \{1, \ldots, i\}, \end{cases} \tag{20}$$

$$E(\mathbf{1}_{B_j}\mathbf{1}_{B_k}) = \begin{cases} \frac{i}{n}\frac{(i-1)}{(n-1)} & \text{if} \quad j \neq k,\ j,\ k \in \{i + 1, \ldots, n - i\} \\ \frac{i}{n} & \text{if} \quad j = k,\ j \in \{i + 1, \ldots, n - i\}, \end{cases} \tag{21}$$

and

$$E(\mathbf{1}_{A_j}\mathbf{1}_{B_k}) = \frac{(n - 2i)}{n}\frac{i}{(n - 1)} \quad \text{if} \;\; j \in \{1, \ldots, i\} \text{ and } k \in \{i + 1, \ldots, n - i\}. \tag{22}$$

Therefore,

$$E(T_{i,n}) = E\left(\sum_{j=1}^{i} \mathbf{1}_{A_j} - \sum_{j=i+1}^{n-i} \mathbf{1}_{B_j}\right) = \frac{i(n-2i)}{n} - \frac{(n-2i)i}{n} = 0,$$

and using Eqs. (20), (21) and (22)

$$\mathrm{Var}(T_{i,n}) = E(T_{i,n}^2) = \frac{2i^2(n-2i)}{n(n-1)}.$$

It is also easy to see that $\mathrm{Cov}(\mathbf{1}_{A_j}, \mathbf{1}_{A_k}) = -(n-2i)(2i)/(n^2(n-1))$ for $j \neq k$, $j, k \in \{1, \ldots, i\}$, $\mathrm{Cov}(\mathbf{1}_{B_j}, \mathbf{1}_{B_k}) = -i(n-i)/(n^2(n-1))$ for $j \neq k$, $j, k \in \{i+1, \ldots, n-i\}$, and $\mathrm{Cov}(\mathbf{1}_{A_j}, \mathbf{1}_{B_k}) = (n-2i)i/(n^2(n-1))$ for $j \in \{1, \ldots, i\}$ and $k \in \{i+1, \ldots, n-i\}$. □

We have found the exact distribution of $T_{i,n}$ defined in Eq. (10) under independence. However, for large values of n we may find its asymptotic distribution under standardization and appropriate conditions on i.

Theorem 3 *Let* $\mathbf{X} = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ *be a random sample of size* n *of a continuous random vector* (X, Y) *of independent random variables in* $\mathbb{R}^2$. *Let* $0 < K < 1/2$ *be a fixed value and let* $i = [Kn]$, *where* $[a]$ *denotes the greatest integer less than or equal to* a. *Let us define the random variables* Z_n, V_n, W_n, *and* $T_{i,n}$ *as in Eqs.* (9) *and* (10) *of Theorem* 2. *If we define*

$$X_n = \frac{T_{i,n}}{\sqrt{\frac{2i^2(n-2i)}{n(n-1)}}}, \tag{23}$$

then X_n *is asymptotically distributed* $N(0, 1)$.

Proof Let $0 < K < 1/2$ be a fixed value and let $i = [Kn]$, where $[a]$ denotes the greatest integer less than or equal to a. Using Eq. (19), we can see that $T_{i,n} = \sum_{i=1}^{n} a(i, R_i)$, where $\{a(i, j)\}$ is the $n \times n$ matrix given by

$$a(j, R_j) = \begin{cases} 1 & \text{if } \ 1 \leq j \leq i \ \text{ and } \ i+1 \leq R_j \leq n-i \\ -1 & \text{if } \ i+1 \leq j \leq n-i \ \text{ and } \ 1 \leq R_j \leq i \\ 0 & \text{otherwise} \end{cases}$$

Therefore, $T_{i,n}$ is a linear rank statistic, and the asymptotic normality follows from standard results of R-estimates, see for example Serfling (1980). □

In order to test (7) we will use the statistic $T_{i,n}$ and its distribution under independence. Let $n \geq 3$ be a fixed integer, and let $i = [n/3]$, where $[n/3]$ denotes the greatest integer less than or equal to $n/3$, then $1 \leq i < n/2$. Assume that $\{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ is a random sample of a continuous random vector (X, Y)

with underlying copula $C_{X,Y}$. If we want to test (7), we can use $T_{i,n}$ as defined in (10), with $i = [n/3]$. If n is not a multiple of 3 we use the extension of C_n to $[0, 1]^2$ given before Definition 1 to define T_n. Observe that if H_0 holds, then $|T_{i,n}|$ should be small. Hence, large values of $|T_{i,n}|$ give evidence against H_0. Of course the distribution of $T_{i,n}$ depends on the copula $C_{X,Y}$; however, we can use the distribution of $T_{i,n}$ under independence, since independent samples give large variability for the values of the statistic $T_{i,n}$, so we take independence as representative of the class of absolutely continuous and symmetric copulas. An idea of whether this could be a good choice might be illustrated by the results in the following section.

4 A power study

In this section we include several simulations in order to check the power of the proposed test against two other well known nonparametric symmetry tests: Hollander's test (1971) and Edwards' continuity correction (1948) applied to Bowker's test (1948), in the trisymmetric case (3 categories). Among Bowker's test (1948) and two of its modifications, Edwards' test (1948) and Wald's modified test as introduced by May and Johnson (2001), we chose Edwards', which from now on we will just refer to as Bowker–Edwards test, since its asymptotic distribution works relatively better compared to the other two, see Krampe and Kuhnt (2007). The proposed test, which from now on we will refer to as the Trisymmetric test, rejects the null hypothesis of symmetry whenever $|T_{i,n}| > k_\alpha$ for an appropiate threshold k_α depending on the test size α.

We present power comparisons for a small sample ($n = 15$) and a larger sample ($n = 150$), with a test size level $\alpha = 0.05$. We chose a small sample $n = 15$ in order to be able to compare the proposed exact Trisymmetric test against Hollander's exact test, since Hollander's test statistic is of extreme computational complexity. For $n = 150$ we use the asymptotic distributions of the Trisymmetric, Hollander and Bowker–Edwards test statistics.

As alternatives to the null hypothesis of symmetry, we used four different families of absolutely continuous asymmetric copulas: the one presented in Sect. 2, which we will refer to as 3-square asymmetric copula; the gluing copula resulting from Clayton family and an asymmetric version of it; an asymmetric version of the Raftery family of copulas; a parametric W-ordinal sum of Π copulas, as in De Baets and De Meyer (2004, 2007), Durante et al. (2008b), Mesiar and Szolgay (2004). Asymmetric versions of Clayton and Raftery families of copulas were obtained by applying Theorem 2.4.4 (1) in Nelsen (2006), and the gluing copula technique is the one introduced in Siburg and Stoimenov (2008). The estimated power of the tests were obtained by 10,000 simulations per asymmetry value, and such power is expected to be an increasing function of the degree of asymmetry, which is measured as $\delta(C) = \sup_{(u,v)\in I^2} |C(u, v) - C(v, u)| \in [0, 1/3]$.

In Fig. 7 we show samples of the 3-square asymmetric copula with different levels of asymmetry, from the highest possible level of $1/3$ (accordingly to Lemma 1) to perfect symmetry. In Fig. 8 we show the power of our test against the power of the exact Hollander's test for $n = 15$ and in Fig. 9 we give the power of our test compared to

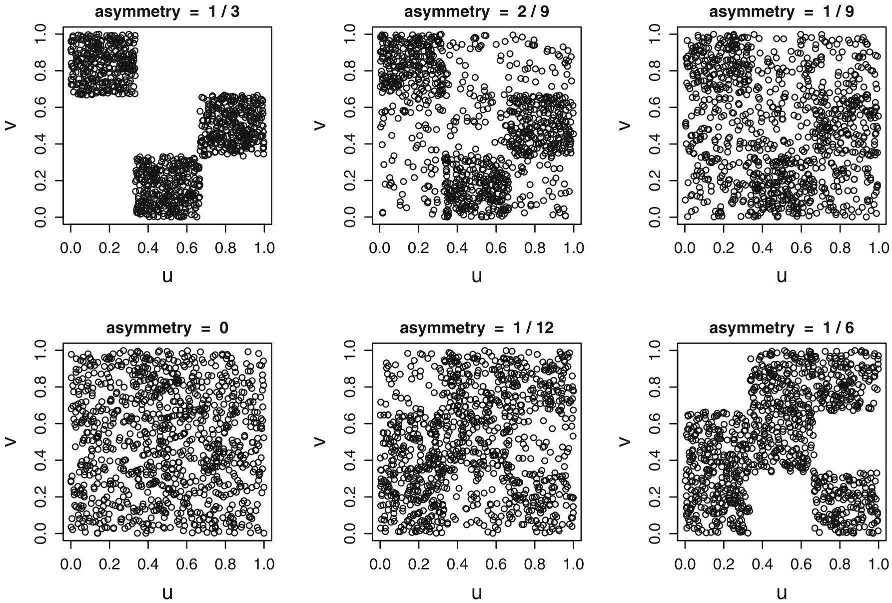


Fig. 7 Samples of the 3-square asymmetric copula under different degrees of asymmetry

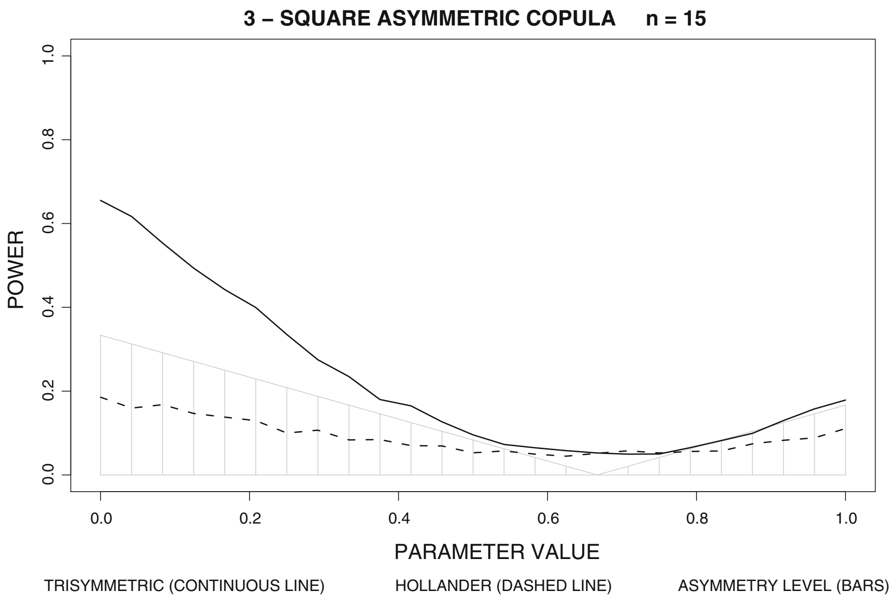


Fig. 8 Power graphs of Trisymmetric and Hollander exact tests for $n = 15$, under the 3-square asymmetric copula

the power of the asymptotic versions of Hollander and Bowker–Edwards for $n = 150$. The vertical bars provide the asymmetry level so that it could be verified if a higher asymmetry level leads to a higher power of the tests. In both cases we notice that our

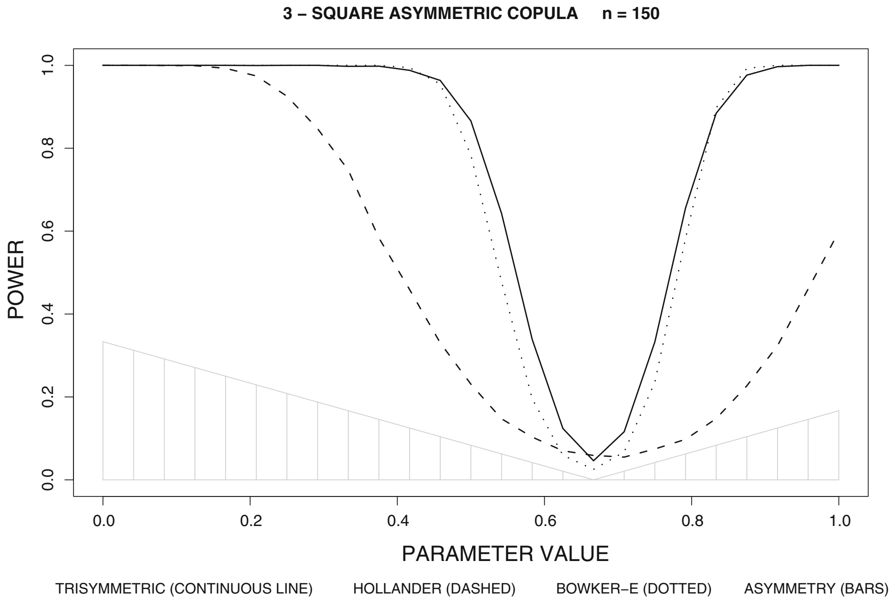


Fig. 9 Power graphs of Trisymmetric, Hollander and Bowker asymptotic tests for $n = 150$, under the 3-square asymmetric copula

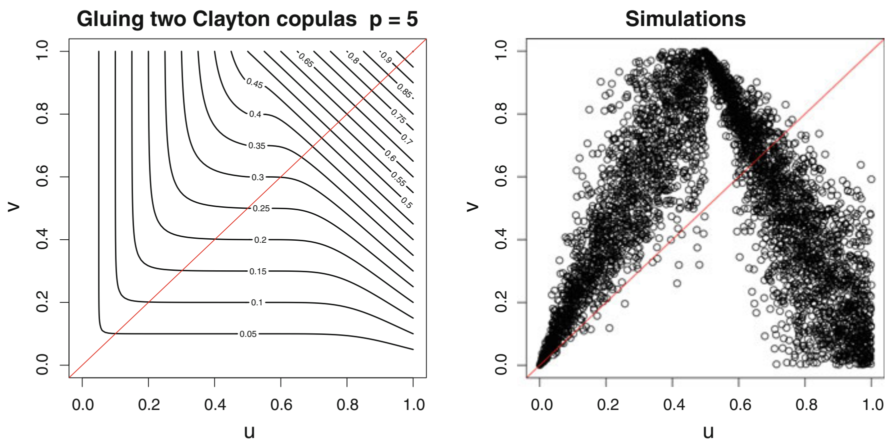


Fig. 10 Gluing a Clayton copula and its asymmetric version: level curves (*left*) and simulations (*right*)

proposed test clearly outperforms Hollander's, and with no significant difference with Bowker–Edwards test.

As a second example, in Fig. 10 we show the level curves of an asymmetric copula which glues a Clayton copula and its asymmetric version. Table 1 shows that for $n = 15$ the power values of our test is better than the exact power of the Hollander test, we do not compare to Bowker–Edwards since $n = 15$ is too small to use the asymptotic distribution, and just a few asymmetry levels are considered since for such

Table 1 Power values of Trisymmetric and Hollander exact tests for $n = 15$ under Gluing Clayton copula and its asymmetric version, for different degrees of asymmetry

Asymmetry	Trisymmetric	Hollander
0.10	0.1621	0.0330
0.15	0.2627	0.0627
0.20	0.3535	0.1142

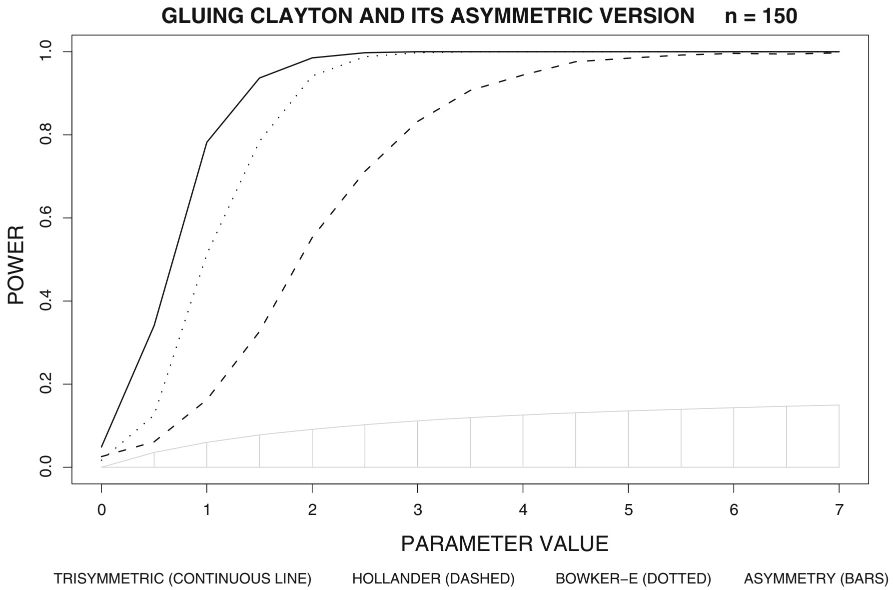


Fig. 11 Power graphs of Trisymmetric, Hollander and Bowker asymptotic tests for $n = 150$, under Gluing Clayton and its asymmetric version

a small sample the graphs of the powers are quite flat. In the case that $n = 150$, Fig. 11 gives the power of our test against the asymptotic versions of Hollander and Bowker–Edwards, with a clear outperformance of our proposed test in comparison to the other two.

As a third example we consider an asymmetric version of the Raftery copula with $p = 0.62$, see Fig. 12 for level curves of this copula. In Table 2 we observe that the power of our test is basically twice the power of the other two tests for $n = 150$. In this particular case we just calculated the power under one asymmetry level (the highest possible) since the asymmetry that may be achieved by this copula is quite low.

As a last example we analize the case of a W-ordinal sum of Π, see Fig. 13 for samples and level curves of these copulas with different asymmetry levels. In Fig. 14, we compare the power of our test to the power of the exact Hollander's test for $n = 15$. Finally in Fig. 15 we compare the power of our test to the asymptotic power of the other two tests for $n = 150$.

As a general remark, it was expected to obtain low power values for a small sample size $n = 15$. In this case, the proposed test clearly performs relatively better than Hollander's exact test under three of the four alternatives under consideration. In the

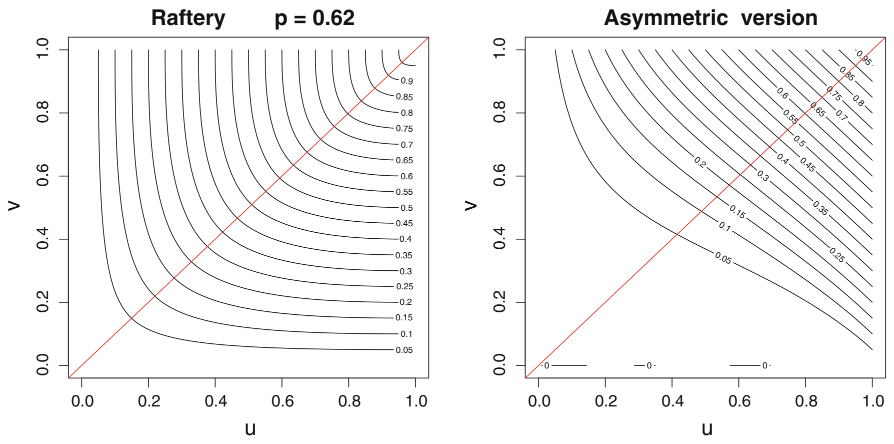


Fig. 12 Raftery copula level curves: symmetric (*left*) and its maximally asymmetric version (*right*)

Table 2 Power of Trisymmetric, Hollander and Bowker asymptotic tests for $n = 150$ under the maximal degree of asymmetry of asymmetric Raftery copula

Asymmetry	Trisymmetric	Hollander	Bowker–Edwards
0.068	0.6215	0.3623	0.3608

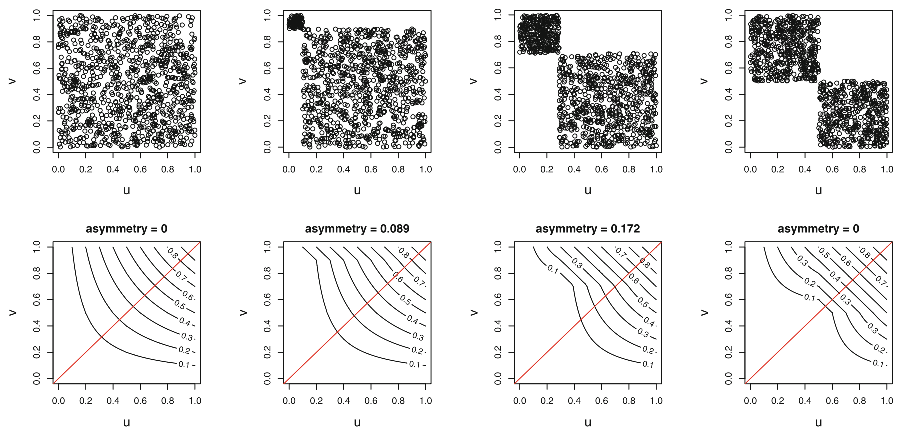


Fig. 13 Samples and level curves of a W-ordinal sum of Π for different degrees of asymmetry

case of the W-ordinal sum of Π copula, eventhough there is a crossing, one may question Hollander's test performance since its power does not reach a maximum near to the point of maximum level of asymmetry, which is approximately 0.172 at a parameter value of 0.39.

For sample size $n = 150$, we have the following comments. Under the 3-square asymmetric copulas, which reaches the maximal level of asymmetry at parameter value of zero, we observe that Hollander's test clearly underperforms the other two tests, with the proposed Trisymmetry test perfoming slightly better than Bowker–Edwards

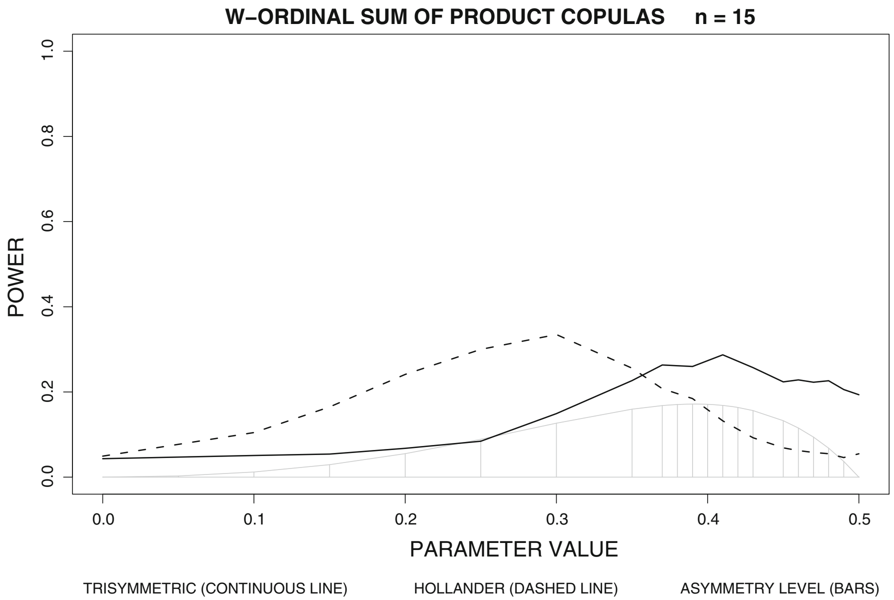


Fig. 14 Power graphs of Trisymmetric and Hollander exact tests for $n = 15$, under a W-ordinal sum of product copulas

almost all the time, but with a crossing at some values of the parameter, see Fig. 9. Under the gluing of Clayton copula and its asymmetric version, whose maximally attainable asymmetry value is $0.2 < 1/3$, the proposed Trisymmetry test performs better that Bowker–Edwards' test, which in turn is better than Hollander's test, see Fig. 11. Under the asymmetric version of Raftery copulas, whose maximally attainable asymmetry value is quite small, approximately 0.068, we observe a better performance of the proposed Trisymmetry test over the other two tests, which show a similar performance, see Table 2. Under the Wordinal sum of Π copulas, we observe crossings of the estimated power functions of the three tests, reaching a maximal power close to where they are expected, the maximal value of asymmetry which is approximately 0.172 at a parameter value of 0.39. In this case we have a not very good power for the proposed Trisymmetry test at a parameter value of 0.5 since at that point the copula becomes symmetric, as it also happens with parameter value equal to zero, and the probability of erroneously rejecting the null hypothesis of symmetry is higher than under the other two tests.

5 Application to real data

According to Erdely and Díaz-Viera (2010) assessment of rock formation permeability is a complicated and challenging problem that plays a key role in oil reservoir modeling, production forecast, and the optimal exploitation management. Generally, permeability evaluation is performed using porosity–permeability relationships obtained by integrated analysis of various petrophysical measurements taken from cores and

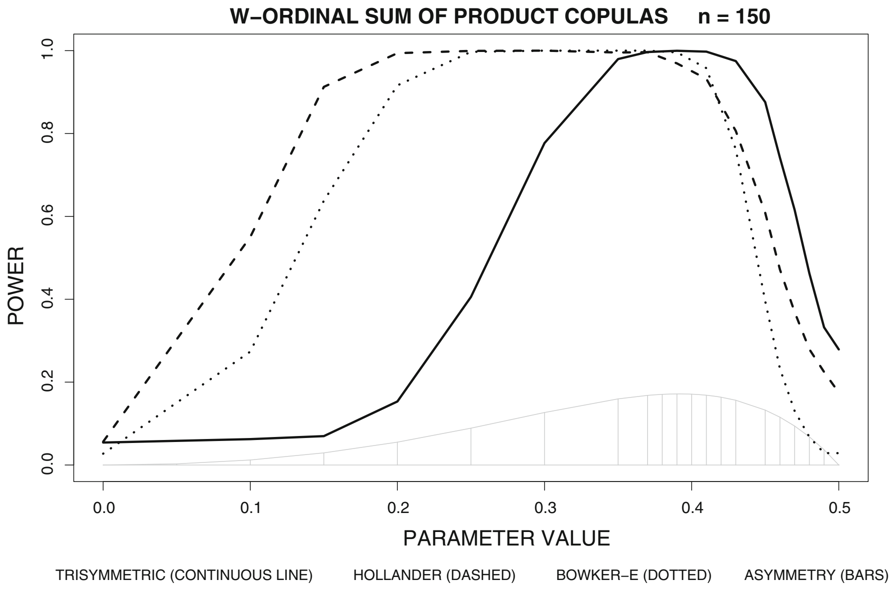


Fig. 15 Power graphs of Trisymmetric, Hollander and Bowker asymptotic tests for $n = 150$, under a W-ordinal sum of product copulas

wireline well logs. Dependence relationships between pairs of petrophysical variables, such as permeability and porosity, are usually nonlinear and complex, and therefore those statistical tools that rely on assumptions of linearity and/or normality and/or existence of moments are commonly not suitable in this case. The use of copulas for modeling petrophysical dependencies is not new, see Díaz-Viera and Casar (2005).

From Kazatchenko et al. (2006) we have bivariate observations of porosity–permeability data, and for subsample 3 in Erdely and Díaz-Viera (2010) the scatterplot of data ranks are shown in Fig. 16.

In terms of choosing a parametric copula, strong evidence against symmetry is challenging since there is not a large catalog of parametric asymmetric copulas as there is indeed for the symmetric case. For the particular data under consideration, it has been possible to transform the data in order to "remove" the asymmetry, by means of Theorem 2.4.4 (1) in Nelsen (2006) and the particular case

$$C_{X,Y}(u, v) = u - C_{X,-Y}(u, 1 - v), \tag{24}$$

with which, fortunately in this case, it was not possible to reject symmetry for the observations of the random vector $(X, -Y)$ from the transformed subsample 3 (denoted as 3T), see Table 3.

In Fig. 17 we show the level curves of the empirical copulas for subsamples 3 and 3T, and by visual inspection it was expected to reject symmetry under subsample 3, but do not reject it under subsample 3T. Hollander's test fails to reject symmetry of subsample 3, and the proposed Trisymmetry test is able to reject symmetry with a better p-value than Bowker–Edwards test.

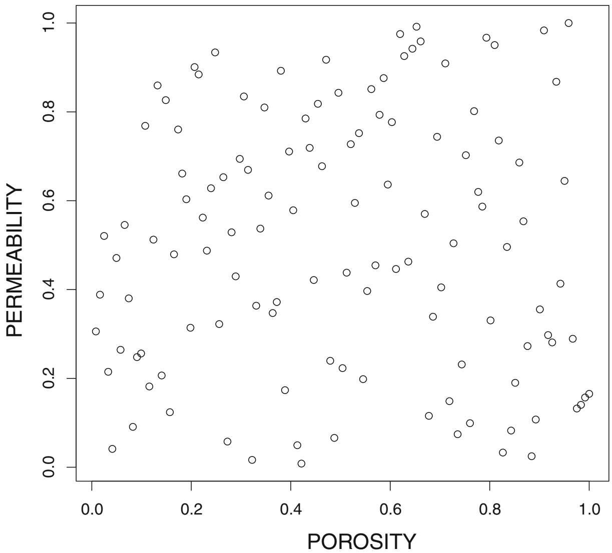


Fig. 16 Scatterplot of porosity–permeability data ranks from subsample 3, rescaled to $[0, 1]^2$

Table 3 *p*-values of symmetry tests applied to subsamples 3 and 3T

Subsample	Trisymmetric test	Hollander test	Bowker–Edwards test
3	0.0014	0.3885	0.0702
3T	0.4040	≈1	0.9181

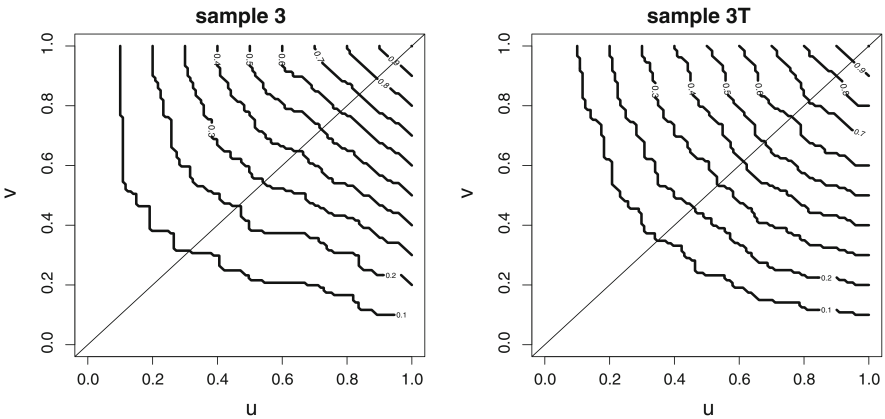


Fig. 17 Level curves of empirical copulas for subsamples 3 and 3T

6 Final remarks

Strictly speaking, the proposed test is a nonparametric trisymmetry test, rather than a symmetry test, but according to the results of the test power study, it shows an adequate performance when used as a nonparametric symmetry test, at least under the alternatives under consideration and compared to two classical tests.

The proposed test makes use of the exact distribution of the test statistic, in contrast with what we commonly find in statistical literature, where the tests usually rely on the asymptotic distribution of the test statistics. An exact test has the advantage that it is useful under small samples. Eventhough asymptotic normality has been proved in Theorem 3, observe that our standardized statistic is discrete and symmetric, but even for small values of n the normal approximation is quite good. For example, for $n = 50$ we performed a large simulation study using Pearson's chi-square to test for normality with 15 classes, and we observed that the proportions of rejections were always close to the α-values of the test. The use of Shapiro–Wilk's test of normality is not recommended in our case due to discretization, but even using this test with the package STATA-10.1 the normality is not rejected for $n = 50$.

The exact test version of Hollander's test makes use of a statistic that is computationally complex, even for small sample sizes, say $n = 20$, it takes a long time to be computed, in comparison to the proposed test whose test statistic is very easily calculated, since we just need the difference of two values of the empirical copula.

To fit or not to fit a symmetric copula? This is an important question that needs to be answered whenever working with real bivariate continuous data. Particularly, if the null hypothesis of symmetry is rejected, then the well known Archimedean family of copulas should be automatically discarded, since non-symmetry implies non-associativity of the underlying copula, and therefore non-archimedeanity. To the best of our knowledge, the problem of designing a nonparametric test for asssociativity (or archimedeanity) remains as an open problem, see Alsina et al. (2003, 2006), so in the meantime a nonparametric test for symmetry might be used as a weak test for such purpose.

Acknowledgments This work was partially supported by Mexico's CONACYT grant 50152.

References

Alsina C, Frank MJ, Schweizer B (2003) Problems on associative functions. Aequ Math 66:128–140

Alsina C, Frank MJ, Schweizer B (2006) Associative functions: triangular norms and copulas. World Scientific Publishing Co, Singapore

Bowker AH (1948) A test for symmetry in contingency tables. J Am Stat Assoc 43:572–574

De Baets B, De Meyer H (2004) From the comparison of orderd lists to the construction of copulas. In: López-Díaz M, Gil MA, Grzegorzewski P, Hyrniewicz O, Lawry L (eds) Soft methodology and random information systems. Springer, Heidelberg, pp 165–172

De Baets B, De Meyer H (2007) Orthogonal grid constructions of copulas. IEEE Trans Fuzzy Syst 15:1053–1062

Deheuvels P (1979) La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. Acad Roy Belg Bull Cl Sci (5) 65(6):274–292

Díaz-Viera M, Casar-González R (2005) Stochastic simulation of complex dependency patterns of petrophysical properties using t-copulas. Proc IAMG'05 GIS Spat Anal 2:749–755

Durante F, Klement EP, Sempi C, Úbeda-Flores M (2008a) Measures of non-exchangeability for bivariate random vectors. Stat Pap doi:10.1007/s00362-008-0153-0

Durante F, Saminger-Platz S, Sarkoci P (2008b) On patchwork techniques for 2-increasing aggregation functions and copulas. In: Dubois D, Lubiano MA, Prade H, Gil MA, Grzegorzewski P, Hryniewicz O (eds) Soft methods for handling variability and imprecision. Springer, Berlin, pp 349–356

Durante F, Sarkoci P, Sempi C (2009) Shuffles of copulas. J Math Anal Appl 352:914–921

Edwards AL (1948) Note on the “correction for continuity” in testing the significance of the difference between correlated proportions. Psychometrika 13:185–187

Erdely A, Díaz-Viera MA (2010) Nonparametric and semiparametric bivariate modeling of petrophysical porosity-permeability well log data. In: Jaworski P, Durante F, Härdle W, Rychlik T (eds) Copula theory and its applications. Springer, Berlin, pp 267–278

Hollander M (1971) A nonparametric test for bivariate symmetry. Biometrika 58:203–212

Kazatchenko E, Markov M, Mousatov A, Parra J (2006) Carbonate microstructure determination by inversion of acoustic and electrical data: application to a South Florida Aquifer. J Appl Geophys 59:1–15

Klement EP, Mesiar R (2006) How non-symmetric can a copula be? Comment Math Univ Carolinae 47:141–148

Krampe A, Kuhnt S (2007) Bowker’s test for symmetry and modifications within the algebraic framework. Comput Stat Data Anal 51:4124–4142

May WL, Johnson WD (2001) Symmetry in square contingency tables: tests fo hypothesis and confidence interval construction. J Biopharm Stat 11:23–33

Mesiar R, Szolgay J (2004) W-ordinals sum of copulas and quasi-copulas. In: Proceedings of MAGIA 2004 conference, Kočovce, Slovak Republic, pp 78–83

Nelsen RB (2007) Extremes of nonexchangeability. Stat Pap 48:329–336

Nelsen RB (2006) An introduction to copulas. Springer, New York

Schweizer B, Sklar A (2005) Probabilistic metric spaces. Dover, New York

Serfling RJ (1980) Approximation theorems of mathematical statistics. John Wiley & Sons, New York

Siburg KF, Stoimenov PA (2008) Gluing copulas. Commun Stat Theory Methods 37:3124–3134