A note on nonparametric estimation of copula-based multivariate extensions of Spearman's rho

Ana Pérez*, Mercedes Prieto-Alaiz†

January 2016

Abstract

Schmid and Schmidt (2007) proposed copula-based nonparametric estimators for some multivariate extensions of Spearman's rho. In this paper, we show that two of those estimators are inappropriate since they can take values out of the parameter space and we discuss alternative proposals.

Keywords: Spearman's rho; Multivariate concordance; Empirical copula, Clayton copula, Gaussian copula.

1 Introduction

There has been several multivariate copula-based measures proposed in the literature to generalize the population bivariate association Spearman's rho; see, for instance, Wolff (1980), Nelsen (1996, 2002), Dolati and Úbeda-Flores (2006), Schmid and Schmidt (2007), Nelsen and Úbeda-Flores (2012) and García *et al.* (2013). See also Joe (1990) for a non-copula-based approach. The problem of estimating such measures has been addressed in Joe (1990) and Schmid and Schmidt (2007). The first author proposes estimators based on ranks and compare their asymptotic efficiency when they are used as test statistics for independence. The second authors suggest plug-in estimators based on empirical copulas and establish their asymptotic normality under rather weak assumptions concerning the copula. García *et al.* (2013) address the estimation problem in the trivariate case.

*Corresponding author. Departamento de Economía Aplicada and IMUVA, Universidad de Valladolid, Spain; Instituto Flores de Lemus, Universidad Carlos III de Madrid, Spain; Avda. Valle Esgueva 6, 47011, Valladolid, Spain; E-mail: perezesp@eaee.uva.es

†Departamento de Economía Aplicada, Universidad de Valladolid, Spain.

The objective of this paper is to show that two of the statistics proposed in Schmid and Schmidt (2007) can not be used as estimators of their population coefficients counterparts, since they could take values out of the parameter space. To overcome this problem, we date back to Joe (1990) and propose alternative nonparametric estimators.

The paper is organized as follows. Section 2 briefly reviews some popular multivariate extensions of the population bivariate association Spearman's rho coefficient. Section 3 focuses on two copula-based multivariate estimators proposed by Schmid and Schmidt (2007) and provides theoretical and empirical evidence of their drawbacks. Section 4 introduces alternative nonparametric estimators and compares, through Monte Carlos experiments, their finite sample performance. Finally, Section 5 concludes the paper with a summary of our main results. Here onwards, we will refer to Schmid and Schmidt (2007) paper as SS07.

2 Multivariate extensions of bivariate Spearman's rho

Let X_1 and X_2 denote two continuous random variables with joint cumulative distribution function F and marginal distribution functions F_1 and F_2, respectively. Let C denote the copula $C : \mathbf{I}^2 \to \mathbf{I}$, where $\mathbf{I} = [0, 1]$, such that $F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$ for all $(x_1, x_2) \in R^2$. Let U_1 and U_2 be uniform random variables defined as the probability integral transformations $U_1 = F_1(X_1)$ and $U_2 = F_2(X_2)$. Then, the copula C is the joint distribution function of (U_1, U_2) and the population bivariate Spearman's rho for X_1 and X_2 can be written in the following two equivalent ways (see Nelsen, 1991):

$$\rho_S = 12 \int_{\mathbf{I}^2} C(u_1, u_2) du_1 du_2 - 3 = 12 \int_{\mathbf{I}^2} u_1 u_2 dC(u_1, u_2) - 3. \tag{1}$$

If we move to a multivariate framework with more than two variables involved, there is not a unique multivariate version of Spearman's ρ_S coefficient. In this section, we focus on two multivariate copula-based versions of ρ_S that were proposed in Wolff (1980) and Nelsen (1996) and were further considered by SS07. Alternative expressions of these two coefficients were introduced in Joe (1990) and will be discussed in Section 4. Other multivariate versions of Spearman's rho, not considered in this paper, have also been proposed; see, for instance, Nelsen (2002), Nelsen and Úbeda-Flores (2012) and García *et al.* (2013).

Let $\mathbf{X} = (X_1, ..., X_d)$ be a d-dimensional continuous random variable with joint distribution function F, marginals $F_1, .., F_d$ and copula $C : \mathbf{I}^d \to \mathbf{I}$ such that $F(x_1, ..., x_d) = C(F_1(x_1), .., F_d(x_d))$ for all $(x_1, ..., x_d) \in R^d$. Let $U_i = F_i(X_i)$ for $i = 1, 2, ..., d$. Then each U_i is uniform on $[0, 1]$ and C is the joint distribution function of $\mathbf{U} = (U_1, ..., U_d)$; see Sklar (1959) for the main results on copulas as the link between joint d-dimensional distribution functions to their one-dimensional margins. If the d variables $X_1, ..., X_d$ were independent, the copula of $\mathbf{X}$ would be the independent copula Π, defined as $\Pi(\mathbf{u}) = \prod_{i=1}^{d} u_i$, for $\mathbf{u} = (u_1, ..., u_d) \in \mathbf{I}^d$. Moreover, the copula C is upper-bounded by

the Fréchet-Hoeffding upper bound M, defined as $M(\mathbf{u}) = \min(u_1, ..., u_d)$. M is a copula that represents maximal dependence, i.e. the case when each of the random variables $X_1, ..., X_d$ is almost surely a strictly increasing function of any of the others.

The first multivariate version of ρ_S that we consider, due to Wolff (1980) and Nelsen (1996), is a generalization of the left-hand side expression in (1) defined as:

$$\rho_d^- = \frac{\int_{\mathbf{I}^d} C(\mathbf{u})d\mathbf{u} - \int_{\mathbf{I}^d} \Pi(\mathbf{u})d\mathbf{u}}{\int_{\mathbf{I}^d} M(\mathbf{u})d\mathbf{u} - \int_{\mathbf{I}^d} \Pi(\mathbf{u})d\mathbf{u}}. \tag{2}$$

The denominator of the expression above represents the maximum value of its own numerator, i.e. its value at the maximal copula $C = M$. Moreover, since $\int_{\mathbf{I}^d} \Pi(\mathbf{u})d\mathbf{u} = 1/2^d$ and $\int_{\mathbf{I}^d} M(\mathbf{u})d\mathbf{u} = 1/(d+1)$ − see Nelsen (1996) − , expression (2) can be written as:

$$\rho_d^- = \frac{(d+1)}{2^d - (d+1)} \left[2^d \int_{\mathbf{I}^d} C(\mathbf{u})d\mathbf{u} - 1 \right]. \tag{3}$$

Following Nelsen (1996), ρ_d^- can be regarded as a multivariate measure of average lower orthant dependence.

The second multivariate version of ρ_S considered in this paper was originally proposed by Nelsen (1996) as a multivariate measure of average upper orthant dependence. This coefficient is a generalization of the right-hand side expression in (1) defined as:

$$\rho_d^+ = \frac{\int_{\mathbf{I}^d} \Pi(\mathbf{u})dC(\mathbf{u}) - \int_{\mathbf{I}^d} \Pi(\mathbf{u})d\mathbf{u}}{\int_{\mathbf{I}^d} \Pi(\mathbf{u})dM(\mathbf{u}) - \int_{\mathbf{I}^d} \Pi(\mathbf{u})d\mathbf{u}}. \tag{4}$$

Again, the denominator of this expression resembles its own numerator evaluated at the maximal copula, i.e. when $C = M$. Moreover, since $\int_{\mathbf{I}^d} \Pi(\mathbf{u})d\mathbf{M}(u) = 1/(d+1)$ − see Nelsen (1996) −, expression (4) can be alternatively written as:

$$\rho_d^+ = \frac{(d+1)}{2^d - (d+1)} \left[2^d \int_{\mathbf{I}^d} \Pi(\mathbf{u})dC(\mathbf{u}) - 1 \right]. \tag{5}$$

When the copula of $\mathbf{X}$ is the upper bound M, both ρ_d^- and ρ_d^+ attain their maximum value, 1, and they become zero when the components of $\mathbf{X}$ are independent, i.e. when $C = \Pi$. A lower bound for both ρ_d^- and ρ_d^+ is $[2^d - (d+1)!]/\{d![2^d - (d+1)]\}$; see Nelsen (1996). For $d = 2$, both ρ_2^- and ρ_2^+ reduce to bivariate Spearman's ρ_S in (1).

The measure ρ_d^- was first defined in Wolff (1980) who denoted it by ρ_d. Unlike, SS07 denote ρ_d^- and ρ_d^+ by ρ_1 and ρ_2, respectively. Moreover, ρ_d^- and ρ_d^+ were already proposed by Joe (1990) as $\overline{\omega}(F)$ and $\omega(F)$, respectively.

3 Some drawbacks of two nonparametric estimators based on empirical copulas

Let $\{(X_{1j}, ..., X_{dj})\}_{j=1,...,n}$ be a sample of n serially independent random vectors from the d-dimensional continuous variable $\mathbf{X} = (X_1, ..., X_d)$ with associated copula C introduced in Section 2. Let R_{ij} be the *rank* of X_{ij} among $\{X_{i1}, ..., X_{in}\}$, with $i = 1, ..., d$

and $j = 1, ..., n$. SS07 estimates the copula C by the empirical copula defined as:

$$\widetilde{C}_n(\mathbf{u}) = \frac{1}{n} \sum_{j=1}^{n} \prod_{i=1}^{d} \mathbf{1}_{\{\widetilde{U}_{ij} \leq u_i\}}, \text{ for } \mathbf{u} = (u_1, ..., u_d) \in \mathbf{I}^d, \tag{6}$$

where $\mathbf{1}_A$ denotes the indicator function on a set A and $\widetilde{U}_{ij} = R_{ij}/n$. Then, they propose estimating the coefficients ρ_d^- and ρ_d^+ defined in Section 2 by replacing the copula C in (3) and (5) with the empirical copula in (6), i.e.:

$$\widetilde{\rho}_d^- = h(d) \left[2^d \int_{\mathbf{I}^d} \widetilde{C}_n(\mathbf{u}) d\mathbf{u} - 1 \right] = h(d) \left[\frac{2^d}{n} \sum_{j=1}^{n} \prod_{i=1}^{d} (1 - \widetilde{U}_{ij}) - 1 \right], \tag{7}$$

$$\widetilde{\rho}_d^+ = h(d) \left[2^d \int_{\mathbf{I}^d} \Pi(\mathbf{u}) d\widetilde{C}_n(\mathbf{u}) - 1 \right] = h(d) \left(\frac{2^d}{n} \sum_{j=1}^{n} \prod_{i=1}^{d} \widetilde{U}_{ij} - 1 \right), \tag{8}$$

where $h(d) = (d+1)/[2^d - (d+1)]$. However, as we will next show, these estimators are inappropriate since they can take values out of the parameter space. For instance, the maximum value of $\widetilde{\rho}_d^+$, that is achieved in the case of perfect dependence, i.e. when $\widetilde{U}_{1j} = \widetilde{U}_{2j} = ... = \widetilde{U}_{dj}$ for each j almost surely, is given by

$$h(d) \left[\frac{2^d}{n} \sum_{j=1}^{n} \left(\frac{j}{n} \right)^d - 1 \right]. \tag{9}$$

Therefore, when $d = 2$, the maximum value of $\widetilde{\rho}_2^+$ becomes $1 + 2(3n+1)/n^2$, which is greater than 1. Moreover, it can also be shown that the following relationship holds:

$$\widetilde{\rho}_2^- = -\frac{12}{n} + \widetilde{\rho}_2^+.$$

Hence, if $\widetilde{\rho}_2^+ = -1$, it will turn out that $\widetilde{\rho}_2^- < -1$, which is an unfeasible value for an estimator of ρ_2^-. Also, if $\widetilde{\rho}_2^- = 1$, it will turn out that $\widetilde{\rho}_2^+ > 1$, which is an unfeasible value for an estimator of ρ_2^+. The following example enhances this feature.

Example 1. The following matrices display a simulated sample of size $n = 5$ from a standard bivariate Normal variable (X_1, X_2) with zero correlation, together with its empirical marginal distribution functions:

$$(x_{ij})_{i=1:2,j=1:5} = \begin{pmatrix} -0.933 & -0.248 \\ -0.370 & -2.072 \\ -0.371 & 1.223 \\ 2.555 & -0.532 \\ 0.152 & -0.125 \end{pmatrix}' \quad (\widetilde{U}_{ij})_{i=1:2,j=1:5} = \begin{pmatrix} 0.2 & 0.6 \\ 0.6 & 0.2 \\ 0.4 & 1.0 \\ 1.0 & 0.4 \\ 0.8 & 0.8 \end{pmatrix}'.$$

Applying formulae (7) and (8) to these data, it turns out that $\widetilde{\rho}_2^- = -1.368$ and $\widetilde{\rho}_2^+ = 1.032$, which are both clearly outside the parametric space $[-1, 1]$.

When $d = 3$, both ρ_3^- and ρ_3^+ are bounded to take values in $[-2/3, 1]$. However, it turns out from (9) that the maximum value of $\widetilde{\rho}_3^+$ becomes $1 + 2(n+1)/n^2$, which is

clearly greater than 1. Moreover, our following example illustrates that $\widetilde{\rho}_3^-$ could also take values out of the parameter range.

Example 2. The following matrices display a simulated sample of size $n = 5$ from a standard trivariate Normal variable (X_1, X_2, X_3) where the correlation matrix is the identity matrix, together with its empirical marginal distribution functions:

$$(x_{ij})_{i=1:3,j=1:5} = \begin{pmatrix} 1.138 & -1.058 & 0.109 \\ -0.346 & -1.031 & 0.846 \\ -0.210 & 0.557 & -0.141 \\ -0.084 & 1.483 & -0.679 \\ 1.033 & 0.536 & 0.632 \end{pmatrix}' \quad (\widetilde{U}_{ij})_{i=1:3,j=1:5} = \begin{pmatrix} 1.0 & 0.2 & 0.6 \\ 0.2 & 0.4 & 1.0 \\ 0.4 & 0.8 & 0.4 \\ 0.6 & 1.0 & 0.2 \\ 0.8 & 0.6 & 0.8 \end{pmatrix}'.$$

Applying formulae (7) and (8) to these data, we obtain $\widetilde{\rho}_3^- = -0.859$ and $\widetilde{\rho}_3^+ = 0.331$, the former being less than the theoretical parametric lower bound $-2/3$.

In higher dimensions, the upper bound for $\widetilde{\rho}_d^+$ in (9) can be evaluated using the formula 0.121 in Gradshteyn and Ryzhik (1994) and similar results would come up. As expected, this bound converges to 1 as $n \to \infty$.

To reinforce the arguments above, we next estimate, via Monte Carlo simulations, the probability that a sample yields a Spearman's rho exceeding the theoretical parameter range for a given copula model. In order to do that we generate samples from the d-dimensional Clayton copula:

$$C(\mathbf{u};\theta) = C(u_1, ..., u_d; \theta) = \left(u_1^{-\theta} + \cdots + u_d^{-\theta} - d + 1\right)^{1/\theta},$$

with $\theta > 0$; see Nelsen (2006, p. 152). This copula is tail asymmetric, exhibiting greater dependence in the lower orthant than in the upper orthant. Following Blumentritt and Schmid (2014), we also consider an elliptical equicorrelated d-dimensional Gaussian copula with correlation matrix $R = \varrho \mathbf{1}_d \mathbf{1}_d' + (1-\varrho)\mathbf{I}_d$, where $-1/(d-1) < \varrho < 1$, $\mathbf{1}_d$ is a unit column vector and $\mathbf{I}_d$ denotes the identity matrix. In both cases, four dimensions are analyzed: $d = \{2, 3, 4, 5\}$.

For the Clayton copula we take parameter values $\theta = \{0.2, 0.5, 1, 2, 5\}$. These yield the following values of bivariate Spearman's rho (computed by numerical integration): $\rho_S = \{0.135, 0.295, 0.479, 0.682, 0.885\}$. For the Gaussian copula, we use the identity $\varrho = 2\sin(\pi\rho_S/6)$ to choose positive values of ϱ that provide in the bivariate case the same values of ρ_S above; see Joe (1997, p. 54). We also allow for negative values of ϱ that fulfill the restriction $\varrho > -1/(d-1)$ for all dimensions. Hence, we take $\varrho = \{-0.2, -0.1, 0.141, 0.308, 0.496, 0.699, 0.894\}$. With these models and parameter values we cover a wide spectrum of possible bivariate and multivariate relationships.

To analyze the influence of the sample size we take $n = \{20, 40, 50, 100, 500\}$. These sample sizes are frequently encountered in applications of copulas to fields like energy, hydrology or macroeconomics; see for instance, Favre *et al.* (2004), Genest and Favre (2007), Granger *et al.* (2006) and Zimmer (2012). Obviously, in other fields like finance, these sample sizes are unusual since thousands of observations are readily available.

For each copula model, parameter value and dimension, we simulate 1000 replicates of size n using the *Copula Package* in R. Then, for each replicate, we compute both $\widetilde{\rho}_d^-$ and $\widetilde{\rho}_d^+$ and we estimate the probability that these exceed the theoretical parameter range as the proportion of replicates where this happens. As expected, the estimated probabilities converge to zero as the sample size increases. Therefore, we focus our discussion on sample sizes $n = \{20, 40, 50\}$.

Figure 1 displays a curve of the estimated probabilities of $\widetilde{\rho}_d^+$ to be outside the theoretical parameter range as a function of the parameter value arranged by copula model (columns) and dimension (rows). In each panel the curves for $n = \{20, 40, 50\}$ are displayed. Several conclusions emerge from this figure. First, the problem of getting a value of $\widetilde{\rho}_d^+$ outside the theoretical parameter range is remarkable in small samples and small dimension settings. Second, the larger the parameter value, i.e., the larger the dependence in the data, the larger the probability that this occurs. For instance, in both copula models, even with samples of size $n = 50$, there is a probability around 50% of getting a value of $\widetilde{\rho}_2^+$ exceeding the parameter space if the parameters take the highest values considered. This probability is still around 25% in the 3−dimensional Gaussian copula. Third, for a given sample size n, the estimated probability that $\widetilde{\rho}_d^+$ exceeds the parameter range decreases as the dimension d increases. Finally, as expected, such probability also decreases as the sample size increases.

Regarding $\widetilde{\rho}_d^-$, the probability of this estimator to be outside the theoretical range is zero in all cases considered. However, we have checked that in the bidimensional case, $\widetilde{\rho}_2^-$ could exceed the parameter range in models with stronger negative values of parameter ϱ and small samples. The results are available upon request.

INSERT FIGURE 1 AROUND HERE

4 Alternative nonparametric estimators of multivariate dependence

Joe (1990) already proposed an estimator of ρ_d^+ based on ranks. In this section, we work out an alternative expression of this estimator and propose an estimator of ρ_d^-. We also discuss modified-SS07 alternatives based on using the so-called pseudo-observations, $U_{ij}^* = R_{ij}/(n+1)$, rather than $\widetilde{U}_{ij} = R_{ij}/n$, and we compare them both analytically and by simulations.

4.1 Definition of the estimators

The multivariate coefficient of concordance ρ_d^+ introduced in section 2 was already proposed by Joe (1990) as a scaled expected value of $F_1(X_1)\cdots F_d(X_d)$, namely:[1]

$$\rho_d^+ = \frac{E[F_1(X_1)\cdots F_d(X_d)] - c_1}{c_2}, \tag{10}$$

[1] Note that $E[F_1(X_1)\cdots F_d(X_d)] = E(U_1\cdots U_d) = \int_{\mathbf{I}^d} \Pi(\mathbf{u}) dC(\mathbf{u})$.

where $c_1 = E[F_1(X_1)] \cdots E[F_d(X_d)] = E(U_1) \cdots E(U_d) = (1/2)^d$ and c_2 stands for the value of the numerator in (10) when the joint distribution of $(X_1, ..., X_d)$ is the upper Fréchet-Hoeffding bound, i.e. when $F_1(X_1) = ... = F_d(X_d)$ with probability one. Hence, $c_2 = 1/(d+1) - 1/2^d$. The sample version of (10) in Joe (1990) is:

$$\widehat{\rho}_d^+ = \frac{\frac{1}{n}\sum_{j=1}^{n}\prod_{i=1}^{d} R_{ij} - \left(\frac{n+1}{2}\right)^d}{\frac{1}{n}\sum_{j=1}^{n} j^d - \left(\frac{n+1}{2}\right)^d}. \tag{11}$$

The motivation behind this estimator is based on estimating in (10) the three parameters involved, namely the expectation, say $c_0 = E[F_1(X_1) \cdots F_d(X_d)]$, and the parameters c_1 and c_2. By contrast, the Schmid and Schmidt's statistic $\widetilde{\rho}_d^+$ defined in (8) only estimates c_0 and keeps the constants c_1 and c_2 as known. The parameter c_0 is estimated by replacing the expectation of the product by the corresponding sample product moment, i.e.,

$$\widehat{c}_0 = \frac{1}{n}\sum_{j=1}^{n}\prod_{i=1}^{d}\widetilde{U}_{ij}, \tag{12}$$

The parameter c_1 in (10) is itself a product of expectations. Hence, when each of these expectations is estimated by its corresponding sample average, the following estimator of c_1 turns out:

$$\widehat{c}_1 = \prod_{i=1}^{d}\left(\frac{1}{n}\sum_{j=1}^{n}\widetilde{U}_{ij}\right). \tag{13}$$

Finally, the parameter c_2 in (10) will be estimated by the corresponding sample version of the numerator of (10) evaluated in the case of perfect dependence, i.e., when the ranks in each dimension coincide. In particular, if we take $\widetilde{U}_{ij} = R_{ij}/n$, as in SS07, the following estimation of c_1 and c_2 will come up:

$$\widehat{c}_1 = \left(\frac{n+1}{2n}\right)^d, \quad \widehat{c}_2 = \frac{1}{n}\sum_{j=1}^{n}\left(\frac{j}{n}\right)^d - \left(\frac{n+1}{2n}\right)^d. \tag{14}$$

Now, putting (12) and (14) back together, the estimator of ρ_d^+ is obtained as:

$$\widehat{\rho}_d^+ = \frac{\widehat{c}_0 - \widehat{c}_1}{\widehat{c}_2} = \frac{\frac{1}{n}\sum_{j=1}^{n}\prod_{i=1}^{d}\widetilde{U}_{ij} - \left(\frac{n+1}{2n}\right)^d}{\frac{1}{n}\sum_{j=1}^{n}\left(\frac{j}{n}\right)^d - \left(\frac{n+1}{2n}\right)^d} \tag{15}$$

Note that (15) collapses to (11) by just multiplying both the numerator and the denominator of the former by n^d. By construction, the maximum value of $\widehat{\rho}_d^+$ is 1.

Joe (1990) suggests another multivariate generalization of Spearman's ρ_S that consists of replacing $F_1, \cdots, F_d$ in (10) by $\overline{F}_1, \cdots, \overline{F}_d$, where the latter are the corresponding survival functions, namely $\overline{F}_i(x_i) = p(X_i > x_i) = 1 - F_i(x_i)$, for $i = 1, ..., d$.

In doing so, the coefficient ρ_d^- in (3) will come up as the scaled expected value of $\overline{F}_1(X_1)\cdots\overline{F}_d(X_d)$, given by:[2]

$$\rho_d^- = \frac{E[\overline{F}_1(X_1)\cdots\overline{F}_d(X_d)] - \overline{c}_1}{\overline{c}_2}, \tag{16}$$

where the parameter $\overline{c}_1$ is now regarded as the product of the expectations of the survival functions, rather than the cumulative distribution functions, that is:

$$\overline{c}_1 = E[\overline{F}_1(X_1)]\cdots E[\overline{F}_d(X_d)] = (1/2)^d,$$

and $\overline{c}_2$ stands for the value of the numerator in (16) when the joint distribution of $(X_1, ..., X_d)$ is the upper Fréchet-Hoeffding bound, i.e. $\overline{c}_2 = 1/(d+1) - 1/2^d$.

From expression (16), it seems clear that the Schmid and Schmidt's statistic $\widetilde{\rho}_d^-$ in (7) consists of only estimating the expectation in (16) while keeping the constants $\overline{c}_1$ and $\overline{c}_2$ as known. However, following the motivation of the estimator $\widehat{\rho}_d^+$ explained before, we suggest an alternative estimator of ρ_d^- based on estimating in (16) both the expectation, that will be denoted by $\overline{c}_0 = E[\overline{F}_1(X_1)\cdots\overline{F}_d(X_d)]$, and the parameters $\overline{c}_1$ and $\overline{c}_2$. In order to do that, let us define $\overline{R}_{ij} = n + 1 - R_{ij}$, as in García *et al.* (2013), and set $\widetilde{\overline{U}}_{ij} = \overline{R}_{ij}/n$. Following the same argument as before, we have:

$$\widehat{\overline{c}}_0 = \frac{1}{n}\sum_{j=1}^{n}\prod_{i=1}^{d}\widetilde{\overline{U}}_{ij}, \quad \widehat{\overline{c}}_1 = \prod_{i=1}^{d}\left(\frac{1}{n}\sum_{j=1}^{n}\widetilde{\overline{U}}_{ij}\right) = \left(\frac{n+1}{2n}\right)^d. \tag{17}$$

Finally, the parameter $\overline{c}_2$ is estimated with the corresponding sample version of the numerator of (16) evaluated in the case of perfect dependence, and we end up with the following estimator of ρ_d^-:

$$\widehat{\rho}_d^- = \frac{\widehat{\overline{c}}_0 - \widehat{\overline{c}}_1}{\widehat{\overline{c}}_2} = \frac{\frac{1}{n}\sum_{j=1}^{n}\prod_{i=1}^{d}\widetilde{\overline{U}}_{ij} - \left(\frac{n+1}{2n}\right)^d}{\frac{1}{n}\sum_{j=1}^{n}\left(\frac{j}{n}\right)^d - \left(\frac{n+1}{2n}\right)^d}. \tag{18}$$

Again, multiplying both the numerator and the denominator of (18) by n^d, an alternative expression of $\widehat{\rho}_d^-$ in terms of ranks is obtained, namely:

$$\widehat{\rho}_d^- = \frac{\frac{1}{n}\sum_{j=1}^{n}\prod_{i=1}^{d}\overline{R}_{ij} - \left(\frac{n+1}{2}\right)^d}{\frac{1}{n}\sum_{j=1}^{n}j^d - \left(\frac{n+1}{2}\right)^d}. \tag{19}$$

By construction, this estimator is bounded not to exceed its maximum value 1.

[2] Recall that $E[\overline{F}_1(X_1)\cdots\overline{F}_d(X_d)] = E[(1-U_1)\cdots(1-U_d)] = \int_{\mathbf{I}^d}\overline{\Pi}(\mathbf{u})dC(\mathbf{u}) = \int_{\mathbf{I}^d}C(\mathbf{u})d\mathbf{u}$, where $\overline{\Pi}(\mathbf{u}) = \prod_{j=1}^{d}(1-u_j)$; see Lemma 3.1. in Dolati and Úbeda-Flores (2006).

Noticeably, for the bidimensional case ($d = 2$), both $\widehat{\rho}_2^+$ and $\widehat{\rho}_2^-$ reduce to the usual sample bivariate Spearman's r_S. By contrast, neither $\widetilde{\rho}_2^+$ nor $\widetilde{\rho}_2^-$ in SS07 coincide with r_S. Moreover, when $d = 3$, the estimators $\widehat{\rho}_3^+$ and $\widehat{\rho}_3^-$ reduce to:

$$\begin{aligned}\widehat{\rho}_3^+ &= \frac{8}{n(n-1)(n+1)^2}\sum_{j=1}^{n} R_{1j}R_{2j}R_{3j} - \frac{n+1}{n-1},\\ \widehat{\rho}_3^- &= \frac{8}{n(n-1)(n+1)^2}\sum_{j=1}^{n} \overline{R}_{1j}\overline{R}_{2j}\overline{R}_{3j} - \frac{n+1}{n-1}.\end{aligned}$$

These estimators appear in García *et al.* (2013) as particular cases of an estimator for the directional $\rho-$coefficients developed by Nelsen and Úbeda-Flores (2012) in trivariate distributions. Its asymptotic distribution can also be found in García *et al.* (2013).

As a final comment, it should be pointed out that, using the following result:

$$\sum_{j=1}^{n} j^d = \frac{n^{d+1}}{d+1} + O(n^d)$$

(see formula 0.121 in Gradshteyn and Ryzhik (1994)), it can be shown that expressions (15) and (18) are asymptotically equivalent to expressions (8) and (7), respectively. Hence, in large samples, the Schmid and Schmidt's statistics $\widetilde{\rho}_d^-$ and $\widetilde{\rho}_d^+$ defined in Section 3 will provide very similar values to the estimators $\widehat{\rho}_d^+$ and $\widehat{\rho}_d^-$ proposed in this section, though the former are not proper estimators while the latter are. However, in small samples they could become quite different, with the former providing even unfeasible values out of the boundaries, as it was highlighted in Section 3.

To complete this subsection, we recall that some authors have proposed modified-SS07 estimators based on using the so-called *pseudo-observations*, $U_{ij}^* = R_{ij}/(n+1)$, instead of $\widetilde{U}_{ij} = R_{ij}/n$, to avoid the problems on the boundary. In particular, if the SS07 statistics $\widetilde{\rho}_d^-$ and $\widetilde{\rho}_d^+$ defined in (7) and (8) were constructed using U_{ij}^* instead of $\widetilde{U}_{ij}$, the estimators used in Blumentritt and Schmid (2014) and Bedo and Ong (2000), respectively, would come up. Let us denote by $\widetilde{\rho}_{d*}^-$ and $\widetilde{\rho}_{d*}^+$ such estimators. Then, it turns out that $\widetilde{\rho}_{2*}^+ = \widetilde{\rho}_{2*}^- = \frac{n-1}{n+1} r_S$ and so, $\widetilde{\rho}_{2*}^+ < 1$ and $\widetilde{\rho}_{2*}^- < 1$. When $d = 3$, it is not difficult to show that the maximum value of both $\widetilde{\rho}_{3*}^+$ and $\widetilde{\rho}_{3*}^-$ is $(n-1)/(n+1)$ and so, $\widetilde{\rho}_{3*}^+ < 1$ and $\widetilde{\rho}_{3*}^- < 1$. Similar results can be easily worked out for higher dimensions, since for a general d, both $\widetilde{\rho}_{d*}^-$ and $\widetilde{\rho}_{d*}^+$ are bounded above by:

$$k(d,n) = \frac{(d+1)}{2^d - (d+1)}\left[\frac{2^d}{n(n+1)^d}\sum_{j=1}^{n} j^d - 1\right].$$

Therefore, it seems that these estimators fail to achieve the maximum value 1 for maximal dependence and take a narrower range of values that they should be.

Furthermore, the following relationship holds between the estimators $\widetilde{\rho}_{d*}^-$ and $\widetilde{\rho}_{d*}^+$ and the estimators $\widehat{\rho}_d^-$ and $\widehat{\rho}_d^+$ introduced before:

$$\widetilde{\rho}_{d*}^+ = k(d,n)\widehat{\rho}_d^+, \qquad \widetilde{\rho}_{d*}^- = k(d,n)\widehat{\rho}_d^-$$

Thus, their bias and mean squared error (mse) fulfill the following identities:

$$\begin{aligned} bias(\widetilde{\rho}_{d*}^{+}) &= k(d,n)bias(\widehat{\rho}_{d}^{+}) + [k(d,n)-1]\rho_{d}^{+}, \\ mse(\widetilde{\rho}_{d*}^{+}) &= [k(d,n)]^2 mse(\widehat{\rho}_{d}^{+}) + 2k(d,n)[k(d,n)-1]bias(\widehat{\rho}_{d}^{+})\rho_{d}^{+} + [k(d,n)-1]^2\rho_{d}^{+2} \end{aligned}$$

Obviously, the same relationships hold between bias and mse of $\widetilde{\rho}_{d*}^{-}$ and $\widehat{\rho}_{d}^{-}$ and the parameter ρ_{d}^{-}. Moreover, since $k(d,n)$ converges to 1 as $n \to \infty$, the estimators $\widetilde{\rho}_{d*}^{\pm}$ and $\widehat{\rho}_{d}^{\pm}$ are aymptotically equivalent. However, the question arises on how these estimators compare in small samples. The equations above reveal that both bias and rmse depend on d, n and the copula model. In next subsection, we conduct a Monte Carlo study to compare the finite sample performance of these estimators.

Finally, it should be emphasized that the estimators $\widehat{\rho}_{d}^{-}$ and $\widehat{\rho}_{d}^{+}$ in (19) and (11) keep the same regardless of whether we use U_{ij}^{*} or $\widetilde{U}_{ij}$. For instance, if we proceed as we did before to work out expression (11), but we put U_{ij}^{*} instead of $\widetilde{U}_{ij}$ in (12) and (13), we get the following:

$$\begin{aligned} \widehat{c}_{0}^{*} &= \frac{1}{n}\sum_{j=1}^{n}\prod_{i=1}^{d}U_{ij}^{*} = \frac{1}{n(n+1)^d}\sum_{j=1}^{n}\prod_{i=1}^{d}R_{ij}, \\ \widehat{c}_{1}^{*} &= \prod_{i=1}^{d}\left(\frac{1}{n}\sum_{j=1}^{n}U_{ij}^{*}\right) = \frac{1}{2^d}, \;\; \widehat{c}_{2}^{*} = \frac{1}{n(n+1)^d}\sum_{j=1}^{n}j^d - \frac{1}{2^d}. \end{aligned}$$

Now, putting these three values back together to make up the estimator $(\widehat{c}_{0}^{*} - \widehat{c}_{1}^{*})/\widehat{c}_{2}^{*}$ we end up with the same definition of $\widehat{\rho}_{d}^{+}$ in (11). A similar argument shows that in order to derive the expression of our estimator $\widehat{\rho}_{d}^{-}$ in (19), it does not matter whether we use $\widetilde{\overline{U}}_{ij} = \overline{R}_{ij}/n$ or $\overline{U}_{ij}^{*} = \overline{R}_{ij}/(n+1)$ when defining the quantities $\widehat{\overline{c}}_0$, $\widehat{\overline{c}}_1$ and $\widehat{\overline{c}}_2$ in (17) and (18), since the final estimator will become the same.

4.2 Finite sample performance: a comparative study

To assess the finite sample performance of the estimators introduced in the previous subsection, we conduct Monte Carlo simulations for the same d-dimensional copulas described in section 3. We consider four dimensions, $d = \{2, 3, 4, 5\}$, and six sample sizes, $n = \{20, 40, 50, 100, 500, 1000\}$. For each copula, each parameter value and each dimension d, we generate 1000 Monte Carlo replicates of size n and for each replicate, we compute the two estimators of the coefficient ρ_{d}^{+} defined in the previous subsection, namely $\widehat{\rho}_{d}^{+}$ and $\widetilde{\rho}_{d*}^{+}$, and the two estimators of the coefficient ρ_{d}^{-}, namely $\widehat{\rho}_{d}^{-}$ and $\widetilde{\rho}_{d*}^{-}$. Finally, for each estimator we compute the mean and rmse over all replicates. Similar Monte Carlo experiments on $\widetilde{\rho}_{d*}^{-}$ with larger sample sizes can be found in Blumentritt (2012) and Blumentritt and Schmid (2014).

Table 1 displays the results on the estimation of ρ_{d}^{+} for the Clayton copula with dimension $d = 3$, parameter values $\theta = \{0.2, 0.5, 1, 2, 5\}$ and four selected sample sizes

$n = \{20, 50, 100, 500\}$[3]. This table also displays, for each simulated model, an aproximated value of the true Spearman's multivariate rho. These values were obtained by numerical integration or by Monte Carlo simulation as the average of its corresponding sample version in (11) across 300 samples of size 500000. Note that for the Clayton copula, not even the bivariate Spearman's rho has an analytical expression as a function of the parameter θ. Table 2 displays similar results for the coefficient ρ_d^-. Finally, Table 3 reports the results from estimating ρ_d^- in the equicorrelated Gaussian copula with dimension $d = 3$, parameter values $\varrho = \{-0.2, -0.1, 0.141, 0.308, 0.496, 0.699, 0.894\}$ and $n = \{20, 50, 100, 500\}$. Note that since this copula is radially symmetric, $\rho_d^- = \rho_d^+$, thus, only the results for $\widehat{\rho}_d^-$ and $\widetilde{\rho}_{d*}^-$ are displayed (the results for $\widehat{\rho}_d^+$ and $\widetilde{\rho}_{d*}^+$ not displayed here, are nearly the same, as expected).

INSERT TABLE 1 AROUND HERE
INSERT TABLE 2 AROUND HERE
INSERT TABLE 3 AROUND HERE

We first comment the results from the Clayton copula. In terms of bias, both estimators $\widehat{\rho}_d^{\pm}$ and $\widetilde{\rho}_{d*}^{\pm}$ tend to underestimate their corresponding true parameters $\rho_d^{\pm}$ but the former always outperforms the latter. In fact, one could expect a maximum relative bias of 10% in $\widehat{\rho}_d^{\pm}$ while the relative bias in $\widetilde{\rho}_{d*}^{\pm}$ could reach 18%. In terms of rmse, there is not a clear dominance of one estimator over the other. Whereas the estimators $\widetilde{\rho}_{d*}^{\pm}$ provide lower rmse than the estimators $\widehat{\rho}_d^{\pm}$ for lower values of θ (low dependence), the behaviour turns the other way round when the value of θ is large (high dependence). As expected, both the bias and rmse tend to reduce as the sample size increases and the differences between both estimators become negligible in large samples. Additionally, we note that both estimators reproduce properly one of the main features of the Clayton copula, namely its asymmetry. Accordingly, for fixed θ and fix n, it always happens that $\widehat{\rho}_d^- > \widehat{\rho}_d^+$ and $\widetilde{\rho}_{d*}^- > \widetilde{\rho}_{d*}^+$ (compare tables 1 and 2)

Regarding the Gaussian copula (see table 3), similar results arise for the positive values of parameter ϱ. Both estimators $\widehat{\rho}_d^-$ and $\widetilde{\rho}_{d*}^-$ underestimate the parameter ρ_d^- but the former always has less bias. Actually, the maximum relative bias of $\widehat{\rho}_d^-$ is around 5% while the relative bias in $\widetilde{\rho}_{d*}^-$ could reach 17%. As reported by Blumentritt and Schmid (2014), the absolute bias of $\widetilde{\rho}_{d*}^-$ increases steadily along with the parameter ϱ. Actually, for the largest positive values of ϱ the bias of $\widetilde{\rho}_{d*}^-$ in small samples is quite important and much larger than that of $\widehat{\rho}_d^+$. In terms of rmse, both $\widehat{\rho}_d^-$ and $\widetilde{\rho}_{d*}^-$ perform very similarly for moderate positive values of ϱ but when ϱ takes the largest value considered, the former dominates the latter in small samples. When $\varrho < 0$, both estimators overestimate the parameter ρ_d^+ and, in general, $\widehat{\rho}_d^+$ outperforms $\widetilde{\rho}_{d*}^-$ in terms of bias. Both estimators have similar rmse, but the rmse of $\widetilde{\rho}_{d*}^-$ is slightly smaller in small samples and with lower correlation ϱ. Moreover, both estimators seem to estimate

[3]The complete simulation results for all dimensions and all sample sizes are not displayed to save space but are available upon request.

with more precision negative parameters than positive ones. Again, both the bias and rmse tend to reduce as the sample size increases and the differences between both estimators become negligible in large samples, as expected.

Noticeably, we have checked that the results for higher dimensions hardly change.

5 Conclusions

This paper shows that two of the multivariate sample versions of the Spearman's rho coefficient proposed in SS07 can not be used as estimators of their population counterparts, since they could take values out of the parameter space. In turn, we propose alternative nonparametric estimators based on the results in Joe (1990) and we compare them, both analytically and by simulations, with some modified-SS07 estimators based on pseudo-observations. We check that, in general, the former outperforms the latter, especially in small samples and in models with higher dependence. Moreover, the latter do not reach the maximum value 1 when there is maximal dependence and take a narrower range of values than they should.

6 Acknowledgments

Financial support from the Spanish Government under project ECO2012-32401 and from Comunidad de Castilla y León under project VA066U13 is gratefully acknowledged by the first author. The second author acknowledges financial support from the Spanish Government under project ECO2012-32178. The usual disclaimers apply.

References

[1] Bedó, J., Ong, C.S., 2000, Multivariate Spearman's rho for aggregating ranks using copulas. Journal of Machine Learning Research 1, 1–48

[2] Blumentritt, T., 2012. On copula density estimation and measures of multivariate association. Publicac Lohmar, Köln.

[3] Blumentritt, T., Schmid, F., 2014. Nonparametric estimation of copula-based measures of multivariate association from contingency tables. Journal of Statistical Computation and Simulation 84, 781-797.

[4] Dolati, A., Úbeda-Flores, M., 2006. On measures of multivariate concordance. Journal of Probability and Statistical Science 4, 147-163.

[5] Favre, A.C., El Adlouni, S., Perreault, L., Thiémonge, N., Bobé, B., 2004. Multivariate hydrological frequency analysis using copulas .Water Resources Research, 40, W01101, Doi:10.1029/2003wr002456

[6] García, J.E., González-López, V.A., Nelsen, R.B., 2013. A new index to measure positive dependence in trivariate distributions. Journal of Multivariate Analysis 115, 481-495.

[7] Genest, C., Favre, A.C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. Journal of Hydrologic Engineering 12, 347-368.

[8] Gradshteyn, I.S., Ryzhik, I.M., 1994. Table of Integrals, Series and Products. Academic Press, San Diego.

[9] Granger, C.W.J., Teräsvirta, T., Patton, A. J., 2006. Common factors in conditional distributions for bivariate time series. Journal of Econometrics 132, 43–57.

[10] Joe, H., 1990. Multivariate concordance. Journal of Multivariate Analysis 35, 12-30.

[11] Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman and Hall, London.

[12] Nelsen, R.B., 1991. Copulas and association, in: Dall'Aglio, G., Kotz, S., Salinetti, G. (Eds.), Advances in Probability Distributions with Given Marginals: Beyond the Copulas. Kluwer, Dordreccht, pp. 51-74.

[13] Nelsen, R.B., 1996. Nonparametric measures of multivariate association, in: Rüschendorf, L., Schweizer,B., Taylor, M.D. (Eds.), Distributions with Fixed Marginals and Related Topics, IMS Lecture Notes-Monograph Series, vol. 28. Hayward, CA, pp. 223-232.

[14] Nelsen, R.B., 2002. Concordance and copulas: A survey, in: Cuadras, C.M., Fortiana, J., Rodriguez-Lallena, J.A. (Eds.), Distributions with Given Marginals and Statistical Modelling. Kluwer, Dordrecht, pp. 169-178.

[15] Nelsen, R.B., 2006. An Introduction to Copulas, 2nd ed. Springer, NY.

[16] Nelsen, R.B., Úbeda-Flores, M., 2012. Directional dependence in multivariate distributions. Ann Inst Stat Math 64, 677-685.

[17] Schmid, F., Schmidt, R., 2007. Multivariate extensions of Spearman's rho and related statistics. Statistical and Probability Letters 77, 407-416.

[18] Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges. Pub. Inst. Statist.Univ. Paris 8, 229-231.

[19] Wolff, E.F., 1980. N-dimensional measures of dependence. Stochastica 4, 175-188.

[20] Zimmer, D.M., 2012. The role of copulas in the housing crisis. The Review of Economics and Statistics 94, 607–620.

Table 1. Monte Carlo results for two estimators of Spearman's multivariate ρ_d^+ based on 1000 samples of size n generated from a d-variate Clayton copula with parameter θ and dimension $d = 3$. For each model the true parameter values of ρ_d^+ are also displayed

Sample size			$n = 20$		$n = 50$		$n = 100$		$n = 500$	
θ	ρ_d^+		mean	rmse	mean	rmse	mean	rmse	mean	rmse
0.2	0.132	$\widehat{\rho}_d^+$	0.119	0.148	0.131	0.096	0.131	0.066	0.131	0.029
		$\widetilde{\rho}_{d*}^+$	0.108	0.135	0.126	0.092	0.128	0.064	0.130	0.029
0.5	0.282	$\widehat{\rho}_d^+$	0.267	0.154	0.278	0.093	0.276	0.064	0.281	0.030
		$\widetilde{\rho}_{d*}^+$	0.242	0.144	0.267	0.091	0.271	0.064	0.280	0.030
1	0.453	$\widehat{\rho}_d^+$	0.437	0.145	0.447	0.090	0.450	0.064	0.451	0.028
		$\widetilde{\rho}_{d*}^+$	0.396	0.142	0.429	0.089	0.441	0.064	0.450	0.028
2	0.648	$\widehat{\rho}_d^+$	0.627	0.125	0.641	0.078	0.645	0.053	0.648	0.023
		$\widetilde{\rho}_{d*}^+$	0.567	0.138	0.616	0.082	0.632	0.054	0.645	0.023
5	0.858	$\widehat{\rho}_d^+$	0.835	0.078	0.849	0.046	0.854	0.031	0.858	0.014
		$\widetilde{\rho}_{d*}^+$	0.756	0.123	0.815	0.062	0.837	0.037	0.854	0.015

Table 2. Monte Carlo results for two estimators of Spearman's multivariate ρ_d^- based on 1000 samples of size n generated from a d-variate Clayton copula with parameter θ and dimension $d = 3$. For each model the true parameter values of ρ_d^- are also displayed

Sample size			$n = 20$		$n = 50$		$n = 100$		$n = 500$	
θ	ρ_d^-		mean	rmse	mean	rmse	mean	rmse	mean	rmse
0.2	0.139	$\widehat{\rho}_d^-$	0.125	0.160	0.136	0.102	0.137	0.070	0.137	0.032
		$\widetilde{\rho}_{d*}^-$	0.113	0.145	0.131	0.098	0.135	0.069	0.136	0.031
0.5	0.308	$\widehat{\rho}_d^-$	0.290	0.172	0.301	0.104	0.300	0.072	0.306	0.034
		$\widetilde{\rho}_{d*}^-$	0.262	0.162	0.289	0.102	0.294	0.072	0.305	0.034
1	0.504	$\widehat{\rho}_d^-$	0.481	0.160	0.493	0.101	0.499	0.071	0.502	0.030
		$\widetilde{\rho}_{d*}^-$	0.436	0.158	0.474	0.101	0.490	0.071	0.500	0.030
2	0.717	$\widehat{\rho}_d^-$	0.686	0.130	0.705	0.080	0.711	0.052	0.716	0.023
		$\widetilde{\rho}_{d*}^-$	0.620	0.150	0.678	0.085	0.697	0.055	0.713	0.023
5	0.911	$\widehat{\rho}_d^-$	0.884	0.065	0.901	0.035	0.905	0.025	0.910	0.010
		$\widetilde{\rho}_{d*}^-$	0.800	0.123	0.865	0.057	0.887	0.033	0.907	0.011

Table 3. Monte Carlo results for two estimators of Spearman's multivariate ρ_d^- based on 1000 samples of size n generated from a d-variate equicorrelated Gaussian copula with parameter ϱ and dimension $d = 3$. For each model the true parameter values of ρ_d^- are also displayed

	Sample size		$n = 20$		$n = 50$		$n = 100$		$n = 500$	
ϱ	ρ_d^-		mean	rmse	mean	rmse	mean	rmse	mean	rmse
-0.2	-0.191	$\widehat{\rho}_d^-$	-0.189	0.104	-0.186	0.068	-0.189	0.047	-0.190	0.021
		$\widetilde{\rho}_{d*}^-$	-0.171	0.096	-0.179	0.066	-0.185	0.046	-0.190	0.021
-0.1	-0.096	$\widehat{\rho}_d^-$	-0.095	0.129	-0.093	0.077	-0.094	0.055	-0.095	0.024
		$\widetilde{\rho}_{d*}^-$	-0.086	0.117	-0.089	0.074	-0.092	0.054	-0.095	0.024
0.141	0.135	$\widehat{\rho}_d^-$	0.130	0.149	0.131	0.096	0.135	0.067	0.134	0.029
		$\widetilde{\rho}_{d*}^-$	0.117	0.136	0.126	0.093	0.132	0.066	0.134	0.029
0.308	0.295	$\widehat{\rho}_d^-$	0.284	0.158	0.289	0.101	0.293	0.067	0.293	0.031
		$\widetilde{\rho}_{d*}^-$	0.257	0.148	0.278	0.098	0.287	0.066	0.292	0.031
0.496	0.479	$\widehat{\rho}_d^-$	0.458	0.155	0.465	0.091	0.472	0.062	0.478	0.029
		$\widetilde{\rho}_{d*}^-$	0.415	0.153	0.446	0.092	0.463	0.063	0.476	0.029
0.699	0.682	$\widehat{\rho}_d^-$	0.652	0.127	0.672	0.072	0.680	0.048	0.681	0.022
		$\widetilde{\rho}_{d*}^-$	0.590	0.144	0.646	0.077	0.667	0.049	0.678	0.022
0.894	0.885	$\widehat{\rho}_d^-$	0.861	0.066	0.875	0.035	0.880	0.025	0.884	0.010
		$\widetilde{\rho}_{d*}^-$	0.779	0.119	0.841	0.055	0.862	0.033	0.880	0.011

Figure 1: Estimated probabilities of $\widetilde{\rho}_d^+$ to be outside the theoretical parameter range as a function of the parameter value arranged by copula model (columns) and dimension (rows). Four dimensions are considered: $d = 2$ (1st row), $d = 3$ (2nd row), $d = 4$ (3rd row) and $d = 5$ (4rth row). In each panel the estimated probabilities for sample sizes $n = \{20, 40, 50\}$ are displayed.