Transactions on Fuzzy Sets and Systems
URL: https://tfss.journals.iau.ir/
Online ISSN: 2821-0131
Vol.1, No.2, (2022), 1-5
DOI: http://doi.org/10.30495/TFSS.2022.1956899.1024

Article Type: Original Research Article (Invited Paper)

A Note on the Maximum Difference Between Schweizer and Wolff's σ and the Absolute Value of Spearman's ρ

Manuel Úbeda-Flores

Abstract. In this note we correct an error on the possible maximum difference between the (measure of dependence) Schweizer and Wolff's σ and the absolute value of the (measure of concordance) Spearman's ρ given in [8]. Moreover, we provide a possible value for that possible, leaving its proof as an open problem.

AMS Subject Classification 2020: 60E05; 62E10
Keywords and Phrases: Copula, Schweizer and Wolff's σ, Spearman's ρ.

1 Introduction

Aggregation functions play an important role in many applications of fuzzy set theory and fuzzy logic, among many other fields (see, e.g., [1, 3]). Copulas —multivariate probability distribution functions with uniform univariate margins on $[0,1]$— are special types of conjunctive aggregation functions, and they are used in aggregation processes because they ensure that the aggregation is stable in the sense that small error inputs correspond to small error outputs.

The importance of copulas in probability and statistics comes from *Sklar's theorem* [9], which states that the joint distribution H of a pair of random variables (X,Y) and the corresponding (univariate) marginal distributions F and G are linked by a copula C in the following manner:

$$H(x,y) = C\left(F(x), G(y)\right) \text{ for all } (x,y) \in [-\infty,\infty]^2.$$

If F and G are continuous, then the copula is unique; otherwise, C is uniquely determined on $(\text{Range}\, F) \times (\text{Range}\, G)$. For a review on copulas, we refer to the monographs [2, 5]

A (bivariate) *copula* is a function $C\colon [0,1]^2 \longrightarrow [0,1]$ which satisfies:

($C1$) the boundary conditions $C(t,0) = C(0,t) = 0$ and $C(t,1) = C(1,t) = t$ for all t in $[0,1]$, and

($C2$) the 2-*increasing property*, i.e., $V_C(R) := C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0$, where $R = [u_1,u_2] \times [v_1,v_2]$ is a rectangle in $[0,1]^2$.

Corresponding author: Manuel Úbeda-Flores, Email: mubeda@ual.es, ORCID: 0000-0002-0675-9646
Received: 18 April 2022; Revised: 9 May 2022; Accepted: 11 May 2022; Published Online: 7 November 2022.

How to cite: M. Úbeda-Flores, A note on the maximum difference between Schweizer and Wolff's σ and the absolute value of Spearman's ρ, *Trans. Fuzzy Sets Syst.*, 1(2) (2022), 1-5.

The fundamental best-possible bounds inequality for the set of copulas is given by the Fréchet-Hoeffding bounds, i.e., for any copula C we have

$$W(u,v) := \max(0, u+v-1) \leq C(u,v) \leq \min(u,v) =: M(u,v) \tag{1}$$

for all $(u,v) \in [0,1]^2$. The hyperbolic paraboloid $z = uv$ —which corresponds to the copula for independence random variables, denoted by Π— sits midway between M and W.

Let $\mathcal{B}([0,1])$ and $\mathcal{B}([0,1]^2)$ denote the Borel σ-algebras in $[0,1]$ and $[0,1]^2$, respectively. A measure μ on $\mathcal{B}([0,1]^2)$ is *doubly stochastic* if $\mu(B \times [0,1]) = \mu([0,1] \times B) = \lambda(B)$ for every $B \in \mathcal{B}([0,1])$, where λ denotes the Lebesgue measure on $[0,1]$ (see [4] for details). Each copula C induces a doubly stochastic measure μ_C by setting $\mu_C(R) = V_C(R)$ for every rectangle $R \subseteq [0,1]^2$ and extending μ_C to $\mathcal{B}([0,1]^2)$. The *support* of a copula C is the complement of the union of all open subsets of $[0,1]^2$ with μ_C-measure zero, and when we refer to "mass" on a set, we mean the value of μ_C for that set.

In 1904, Charles Spearman defined the *Spearman's* ρ coefficient [10], a measure of concordance according to the set of axioms proposed by Scarsini [7]. For a pair of continuous random variables (X,Y) with associated copula C, the population version of this measure is given by

$$\rho(X,Y) = \rho_C = 12 \int_0^1 \int_0^1 [C(u,v) - uv] \, \mathrm{d}u \mathrm{d}v.$$

It represents the difference of the volume formed by the surfaces $z = C(u,v)$ and $z = uv$ on $[0,1]^2$, and where $\rho_W = -1$ and $\rho_M = 1$.

In 1959, A. Rényi proposed a set of desirable axioms for a nonparametric dependence measure for two continuously distributed random variables (X,Y) [6]. Later, those axioms were conveniently modified by Schweizer and Wolff in [8], where the authors introduced a new measure, called the *Schweizer and Wolff's* σ based upon the distance L_1 between the graphs of a copula C and Π, and which, suitably normalized, is given by

$$\sigma(X,Y) = \sigma_C = 12 \int_0^1 \int_0^1 |C(u,v) - uv| \, \mathrm{d}u \mathrm{d}v,$$

where $(X,Y) \sim C$. Note that, in this case, we have $\sigma_M = \sigma_W = 1$.

For any copula C, the quantity $|\rho_C|$ satisfies all the axioms for a measure of dependence except the fact that $\rho_C = 0$ does not necessarily imply that the random variables are independent (note that $\sigma_C = 0$ if, and only if, $C = \Pi$). If the copula C satisfies $C(u,v) \geq uv$ or $C(u,v) \leq uv$ for all $(u,v) \in [0,1]^2$, then we have $\sigma_C = |\rho_C|$; but if this is not the case, σ_C is often a better measure than ρ (see [8] for several examples).

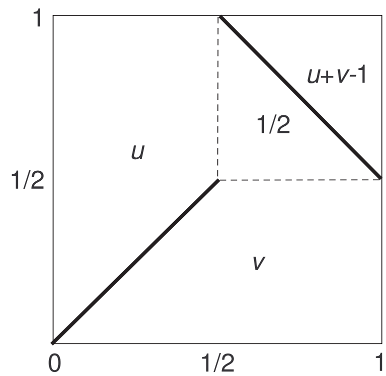
For any copula C, it is clear that $\sigma_C \geq |\rho_C|$. In [8], the authors provide an example for a possible maximum difference between the Schweizer and Wolff's σ and the absolute value of Spearman's ρ, which is ≈ 0.58 (see also [11]); however, this quantity is wrong. In the next section, we correct that error and provide a value (even greater than 0.58) for that possible difference, leaving its proof as an open problem.

2 The example, the correction and the conjecture

In [8] the authors provide the following example of the possible maximum difference between the Schweizer and Wolff's σ and the absolute value of Spearman's ρ.

Example 2.1 ([8]). Let (X,Y) be a pair of continuous random variables such that X is the identity map on $[0,1]$ and Y is defined by

$$Y(w) = \begin{cases} w, & 0 \leq w \leq \dfrac{1}{2} \\ \dfrac{3}{2} - w, & \dfrac{1}{2} < w \leq 1. \end{cases}$$



(a) Support of C (b) $|C(u,v)-uv|$

Figure 1: Support of the copula C and the values of $|C(u,v)-uv|$ in Remark 2.2.

Then $\sigma(X,Y)-|\rho(X,Y)|=3\ln 2-3/2\approx 0.58$.

Remark 2.2. *The difference given in Example* 2.1 is not correct. Note that the copula C, associated with the pair (X,Y), is given by

$$C(u,v)=\begin{cases}\max(1/2,u+v-1), & (u,v)\in[1/2,1]^2,\\ M(u,v), & \text{otherwise.}\end{cases}$$

C is the copula whose mass is spread in two line segments, one joining the points $(0,0)$ to $(1/2,1/2)$ and the other the points $(1/2,1)$ to $(1,1/2)$. Figure 1 shows the support of the copula C and the values of $|C(u,v)-uv|$ for all $(u,v)\in[0,1]^2$.

Then, after some elementary algebra, we obtain $\sigma_C=3\ln 2-5/4\approx 0.83$ *and* $\rho_C=0.75$ *—in* [11] *it appears* 0.25*—; whence* $\sigma_C-|\rho_C|\approx 0.08$.

In the next example, we propose a possible maximum difference $\sigma_C-|\rho_C|$ for a given copula C, even greater than the wrong value 0.58 in Example 2.1.

Example 2.3. Let $0\le\theta\le 1$, and let C_θ be the copula given by

$$C_\theta(u,v)=\begin{cases}\max(0,u+v-\theta), & (u,v)\in[0,\theta]^2,\\ \max(\theta,u+v-1), & (u,v)\in[\theta,1]^2,\\ M(u,v), & \text{otherwise.}\end{cases}$$

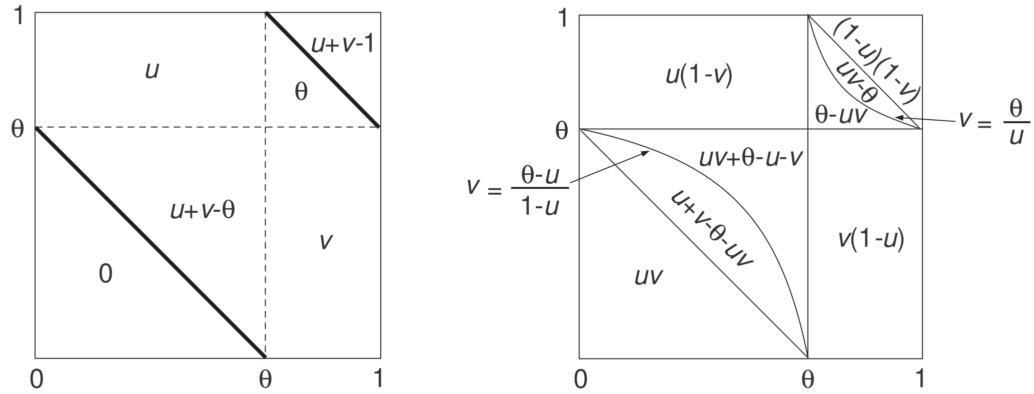
C_θ is the copula whose mass is spread in two line segments, one joining the points $(0,\theta)$ to $(\theta,0)$ and the other the points $(\theta,1)$ to $(1,\theta)$. Figure 2 shows the support of the copula C_θ and the values of $|C_\theta(u,v)-uv|$ for all $(u,v)\in[0,1]^2$. After some algebra, we obtain

$$\sigma_{C_\theta}=1-18\theta(1-\theta)-12\theta^2\ln\theta-12(1-\theta)^2\ln(1-\theta)$$

and

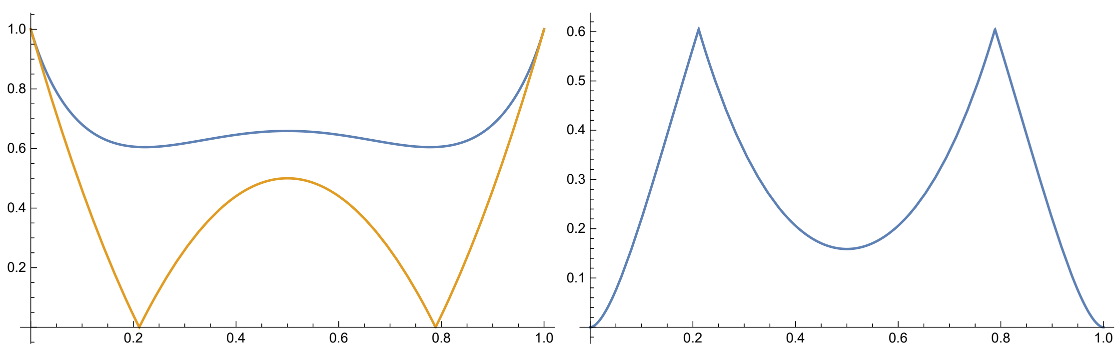
$$\rho_{C_\theta}=-1+6\theta(1-\theta).$$

Figure 3 shows the graphs of σ_{C_θ}, $|\rho_{C_\theta}|$ and $\sigma_{C_\theta}-|\rho_{C_\theta}|$. The maximum of the function $\sigma_{C_\theta}-|\rho_{C_\theta}|$ is reached at the points $\theta_1=1/2-\sqrt{3}/6\approx 0.21$ and $\theta_2=1/2+\sqrt{3}/6\approx 0.79$, for which $\sigma_{C_{\theta_1}}=\sigma_{C_{\theta_2}}\approx 0.60496$ and $\rho_{C_{\theta_1}}=\rho_{C_{\theta_2}}=0$, whence $\sigma_{C_{\theta_1}}-\left|\rho_{C_{\theta_1}}\right|\approx 0.60496$.



(a) Support of C_θ (b) $|C_\theta(u,v) - uv|$

Figure 2: Support of the copula C_θ and the values of $|C_\theta(u,v) - uv|$ in Example 2.3.



(a) σ_{C_θ} (blue) and $|\rho_{C_\theta}|$ (red) (b) $\sigma_{C_\theta} - |\rho_{C_\theta}|$

Figure 3: Support of the copula C_θ and the values of $|C_\theta(u,v) - uv|$ in Example 2.3.

It remains as an open problem to check if, for any copula C, the maximum difference of $\sigma_C - |\rho_C|$ is the quantity given in Example 2.3.

Acknowledgements: The author acknowledges the support of the program FEDER-Andalucía 2014-2020 under research project UAL2020-AGR-B1783.

Conflict of Interest: The author declares that there are no conflict of interest.

References

[1] G. Beliakov, A. Pradera and T. Calvo, Aggregation Functions: A Guide for Practitioners, *Studies in Fuzziness and Soft, Springer, Berlin*, (2007).

[2] F. Durante and C. Sempi, Principles of Copula Theory, *Chapman & Hall/CRC, Boca Raton*, (2016).

[3] M. Grabisch, J. L. Marichal, R. Mesiar and E. Pap, Aggregation Functions, *Encyclopedia of Mathematics and its Applications, 127, Cambridge University Press, Cambridge*, (2009).

[4] P. R. Halmos, Measure Theory, *Springer, New York*, (1974).

[5] R. B. Nelsen, An Introduction to Copulas, Second Edition. *Springer, New York*, (2006).

[6] A. Rényi, On measures of dependence, *Acta Math. Acad. Sci. Hungar.*, 10 (1959), 441-451.

[7] M. Scarsini, On measures of concordance, *Stochastica*, 8 (1984), 201-218.

[8] B. Schweizer and E. F. Wolff, On nonparametric measures of dependence for random variables, *Ann. Statist.*, 9 (1981), 879-885.

[9] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris*, 8 (1959), 229-231.

[10] C. Spearman, The proof and measurement of association between two things, *Am. J. Psychol.*, 15 (1904), 72-101.

[11] E. F. Wolff, Measures of dependence derived from copulas, *PhD Thesis, University of Massachussetts, Amherst, USA*, (1977).

Manuel Úbeda-Flores
Department of Mathematics
University of Almería
Almería, Spain
E-mail: mubeda@ual.es