COMMUNICATIONS IN STATISTICS
Theory and Methods
Vol. 33, No. 10, pp. 2299–2305, 2004

A Comparison of Bounds on Sets of Joint Distribution Functions Derived from Various Measures of Association

Roger B. Nelsen[1,*] **and Manuel Úbeda-Flores**[2]

[1]Department of Mathematical Sciences, Lewis & Clark College, Portland, Oregon, USA
[2]Departamento de Estadística y Matemática Aplicada, Universidad de Almería, Spain

ABSTRACT

We find pointwise best-possible bounds on the bivariate distribution function of continuous random variables with given margins and a given value of the medial correlation coefficient, and compare those bounds to those obtained from a given value of Kendall's tau and Spearman's rho.

Key Words: Blomqvist's beta; Copulas; Kendall's tau; Medial correlation coefficient; Spearman's rho.

---

*Correspondence: Roger B. Nelsen, Department of Mathematical Sciences, Lewis and Clark College, Portland, OR 97219, USA; E-mail: nelsen@lclark.edu.

DOI: 10.1081/STA-200031367 0361-0926 (Print); 1532-415X (Online)

www.dekker.com

1. INTRODUCTION

This article shows that for finding pointwise best-possible bounds on sets of joint distribution functions with given continuous margins and the population version of a nonparametric measure of association, the medial correlation coefficient (also known as Blomqvist's beta) outperforms both Kendall's tau and Spearman's rho for moderate values of these coefficients.

In a previous article Nelsen et al. (2001), the authors (with J. J. Quesada Molina and J. A. Rodríguez Lallena) illustrated a procedure for finding pointwise best-possible bounds on sets of joint distribution functions with given continuous margins and a given value of the population version of a measure of association, such as Kendall's tau or Spearman's rho. The bounds attained are readily evaluated, and hence can be compared (see Sec. 3 below). Before doing so, we use the procedure from Nelsen et al. (2001) to find the bounds on the set of joint distribution functions with given continuous margins and a given value of the population version of the medial correlation coefficient.

As is often the case when dealing with bivariate distribution functions, the use of copulas simplifies matters. A copula is a function $C : \mathbf{I}^2 \to \mathbf{I} = [0, 1]$ that satisfies the boundary conditions

$$C(t, 0) = C(0, t) = 0 \quad \text{and} \quad C(t, 1) = C(1, t)t,\ \ t \in \mathbf{I}, \tag{1}$$

and the two-increasing property

$$C(b, d) - C(a, d) - C(b, c) + C(a, c) \geq 0 \tag{2}$$

for all a, b, c, d in $\mathbf{I}$ such that $a \leq b$ and $c \leq d$. Equivalently, a copula is the restriction to the unit square of a bivariate distribution function whose margins are uniform on $\mathbf{I}$. Recall from Sklar's theorem that any bivariate distribution function H with marginal distribution functions F and G can be written as $H(x, y) = C(F(x),\ G(y))$, where C is a copula. Finally, each copula C satisfies the Fréchet–Hoeffding inequality $W(u, v) = \max(0, u + v - 1) \leq C(u, v) \leq \min(u, v) = M(u, v)$ for u, v in $\mathbf{I}$; furthermore, the Fréchet–Hoeffding bounds W and M are themselves copulas. For further details, see Nelsen (1999).

2. BOUNDS WHEN THE MEDIAL CORRELATION COEFFICIENT IS KNOWN

The population version of the medial correlation coefficient for a pair X,Y of continuous random variables, which we denote as $\beta_{X,Y}$, was first discussed by Blomqvist (1950). If $\tilde{x}$ and $\tilde{y}$ denote medians of X and Y, respectively, then

$$\beta_{X,Y} = P[(X-\tilde{x})(Y-\tilde{y}) > 0] - P[(X-\tilde{x})(Y-\tilde{y}) < 0].$$

Note that $-1 \leq \beta_{X,Y} \leq +1$ and that the bounds are sharp. When H denotes the joint distribution function of X and Y, it readily follows that

$$\beta_{X,Y} = 4H(\tilde{x},\tilde{y}) - 1.$$

Letting C denote the copula of X and Y, we have $H(\tilde{x},\tilde{y}) = C(F(\tilde{x}), G(\tilde{y})) = C(1/2, 1/2)$, and hence,

$$\beta_{X,Y} = \beta(C) = 4C(1/2, 1/2) - 1.$$

For any t in $[-1, 1]$, let $\mathbf{B}_t$ denote the set of copulas with a common value t of the medial correlation coefficient; that is,

$$\mathbf{B}_t = \{C \mid C\text{is a copula}, \beta(C) = t\}.$$

Let $\underline{B}_t$ and $\overline{B}_t$ denote, respectively, the pointwise infimum and supremum of $\mathbf{B}_t$, i.e., for each (u, v) in $\mathbf{I}^2$,

$$\begin{aligned} &\underline{B}_t(u,v) = \inf\{C(u,v) \mid C \in \mathbf{B}_t\} \quad \text{and} \\ &\overline{B}_t(u,v) = \sup\{C(u,v) \mid C \in \mathbf{B}_t\}. \end{aligned} \tag{3}$$

The bounds $\underline{B}_t$ and $\overline{B}_t$ for $\mathbf{B}_t$ are given in Theorem 1.

Theorem 1. *Let* $\underline{B}_t$ *and* $\overline{B}_t$ *denote the pointwise infimum and supremum* (3) *of* $\mathbf{B}_t$, *for* t *in* $[-1,1]$. *Then for any* (u, v) *in* $\mathbf{I}^2$,

$$\underline{B}_t(u,v) = \max\{0, u+v-1, (t+1)/4 - ((1/2)-u)^+ - ((1/2)-v)^+\} \tag{4}$$

and

$$\overline{B}_t(u,v) = \min\{u, v, (t+1)/4 + (u-(1/2))^+ + (v-(1/2))^+\}, \qquad (5)$$

where $x^+ = \max(x,0)$. *Hence, if* X *and* Y *are continuous random variables with joint distribution function* H *and marginal distribution functions* F *and* G, *respectively, and such that* $\beta_{X,Y} = t$, *then the best-possible bounds for* H *are*

$$\underline{B}_t(F(x),G(y)) \le H(x,y) \le \overline{B}_t(F(x),G(y)) \qquad (6)$$

for all (x,y) *in* $(-\infty,\infty)^2$.

Proof. Let $C \in \mathbf{B}_t$. Then, for all (u,v) in $\mathbf{I}^2$, the defining properties (1) and (2) for copulas readily yield the inequalities $-((1/2)-u)^+ \le C(u,v) - C(1/2,v) \le (u-(1/2))^+$ and $-((1/2)-v)^+ \le C(1/2,v) - C(1/2,1/2) \le (v-(1/2))^+$; hence, $(t+1)/4 - ((1/2)-u)^+ - ((1/2)-v)^+ \le C(u,v) \le (t+1)/4 + (u-(1/2))^+ + (v-(1/2))^+$. Furthermore, $W(u,v) \le C(u,v) \le M(u,v)$; thus, $\underline{B}_t(u,v) \le C(u,v) \le \overline{B}_t(u,v)$, where $\underline{B}_t$ and $\overline{B}_t$ satisfy (4) and (5), respectively. However, functions on $\mathbf{I}^2$ of the form $\max\{0, u+v-1, \theta-(a-u)^+-(b-v)^+\}$ and $\min[u, v, \theta+(u-a)^+ + (v-b)^+\}$ are copulas for any (a,b) in $\mathbf{I}^2$ and θ in $[W(a,b), M(a,b)]$ (Nelsen, 1999, Theorem 3.2.2), so $\underline{B}_t$ and $\overline{B}_t$ are copulas. Because $\beta(\underline{B}_t) = \beta\overline{B}_t) = t, \underline{B}_t$ and $\overline{B}_t$ are both in $\mathbf{B}_t$; hence, $\underline{B}_t$ and $\overline{B}_t$ are the pointwise best-possible bounds for $\mathbf{B}_t$.

Because $\underline{B}_t$ and $\overline{B}_t$ are copulas, the bounds for H in (6) are distribution functions. In the following corollary, whose proof is straightforward, we present some additional facts about the bounds $\underline{B}_t$ and $\overline{B}_t$.

Corollary 2. *Let* $\underline{B}_t$ *and* $\overline{B}_t$ *be as in Theorem* 1. *Then,*

(a) $\underline{B}_t$ *and* $\overline{B}_t$ *are continuous and nondecreasing in* t.
(b) $\underline{B}_t = W$ *if and only if* $t = -1$; $\overline{B}_t = M$ *if and only if* $t = +1$.
(c) $\underline{B}_t(u,v) = u - \overline{B}_{-t}(u, 1-v) = v - \overline{B}_{-t}(1-u, v)$, *and similarly for* $\overline{B}_t$.
(d) *Both* $\underline{B}_t$ *and* $\overline{B}_t$ *are radially symmetric, that is,* $\underline{B}_t(u,v) = u + v - 1 + \underline{B}_t(1-u, 1-v)$, *and similarly for* $\overline{B}_t$.

3. A COMPARISON OF THE BOUNDS

Bounds on sets of copulas with a common value of the population versions of the measures of association known as Kendall's tau (τ) and

Spearman's rho (ρ) also exist. Analogous to (3), we let $\underline{T}_t$, $\overline{T}_t$, $\underline{P}_t$, and $\overline{P}_t$ denote the pointwise best-possible lower and upper bounds on $\mathbf{T}_t = \{C \mid C \text{ is a copula, } \tau(C) = t\}$ and $\mathbf{P}_t = \{C \mid C \text{ is a copula, } \rho(C) = t\}$, respectively. These bounds are given explicitly in Nelsen et al. (2001), and like $\underline{B}_t$ and $\overline{B}_t$, are copulas.

Which of the coefficients, β, τ, or ρ, is more effective in narrowing the Fréchet–Hoeffding bounds? To measure the effectiveness of the coefficients for this purpose, we use the function

$$m_\alpha(t) = 1 - 6 \iint_{\mathbf{I}^2} \left[\overline{A}_t(u, v) - \underline{A}_t(u, v)\right] du\, dv, \tag{7}$$

where α denotes a measure of association such as $\beta, \tau,$ or ρ, and $\underline{A}_t$ and $\overline{A}_t$ are the bounds on the set $\mathbf{A}_t = \{C \mid C \text{ is a copula, } \alpha(C) = t\}$. The double integral in (7) represents the volume between the surfaces $z = \underline{A}_t(u, v)$ and $z = \overline{A}_t(u, v)$ in $\mathbf{I}^3$, and m_α is scaled so $m_\alpha(t) = 0$ when there is no improvement in the bounds (i.e., $\underline{A}_t = W$ and $\overline{A}_t = M$), and $m_\alpha(t) = 1$ when the bounds coincide.

In the first four columns of Table 1, we present the values of $m_\beta(t)$, $m_\tau(t)$, and $m_\rho(t)$ for $|t| \in [0, 1]$ (note that $m_\alpha(-t) = m_\alpha(t)$ for $\alpha = \beta, \tau, \text{or } \rho$). Although $m_\beta(t)$ can be computed explicitly $[m_\beta(t) = 3(3t^2 + 1)/16]$, $m_\tau(t)$, and $m_\rho(t)$ must be computed numerically. All results have been rounded to four places.

Comparing column (b) with columns (c) and (d), we see that the medial correlation coefficient is dramatically better than either Kendall's tau

Table 1. A comparison $m_\alpha(t)$ for $\alpha = \beta, \tau$, and ρ.

(a) $\lvert t\rvert$	(b) $m_\beta(t)$	(c) $m_\tau(t)$	(d) $m_\rho(t)$	(e) $m_\rho(t^*)$
0	0.1875	0	0.0295	0.0295
0.1	0.1931	0.0013	0.0327	0.0367
0.2	0.2100	0.0078	0.0425	0.0588
0.3	0.2381	0.0228	0.0596	0.0971
0.4	0.2775	0.0495	0.0851	0.1535
0.5	0.3281	0.0922	0.1209	0.2308
0.6	0.3900	0.1562	0.1700	0.3332
0.7	0.4631	0.2497	0.2386	0.4646
0.8	0.5475	0.3860	0.3389	0.6271
0.9	0.6431	0.5929	0.5019	0.8175
1	0.7500	1	1	1

Table 2. A second comparison of β, τ, and ρ.

$\lvert t\rvert$	$m_\beta(t)$	$m_\tau(8t/9)$	$m_\rho(4t/3)$
0	0.1875	0	0.0295
0.05	0.1889	0.0002	0.0309
0.1	0.1931	0.0010	0.0352
0.15	0.2002	0.0028	0.0425
0.2	0.2100	0.0058	0.0531
0.25	0.2227	0.0103	0.0671

or Spearman's rho for values of $|t| \leq 0.9$. Furthermore, $m_\rho(t)$ is greater than $m_\tau(t)$ for $|t| \leq 0.6$, but the inequality is reversed for $|t| \geq 0.7$. However, this comparison is specious because the three coefficients are rarely equal for a given copula. For example, when C is the copula associated with a standard bivariate normal distribution, $\tau(C) = \beta(C)$; however, $\rho(C) = (6/\pi)\arcsin[(1/2)\sin(\pi\beta(C)/2)]$ (see Kruskal, 1958). If we set $t^* = (6/\pi)\arcsin[(1/2)\sin(\pi t/2)]$ and compare columns (b), (c), and (e) of Table 1, we see that when the dependence structure is described by a normal copula, Spearman's rho is more effective than Kendall's tau in narrowing the bounds for all t, but better than the medial correlation coefficient only for $|t| \geq 0.7$.

Other families of copulas yield different relationships among β, τ, and ρ. Consider, for example, the Farlie–Gumbel–Morgenstern (FGM) family of copulas: for θ in $[-1, 1]$ and all (u, v) in $\mathbf{I}^2$, $C_\theta(u, v) = uv + \theta uv(1-u)(1-v)$. For members of this family, $\beta = \theta/4, \tau = 2\theta/9$, and $\rho = \theta/3$, so $\tau = 8\beta/9$ and $\rho = 4\beta/3$. To compare the effectiveness of β, τ, and ρ for copulas with this relationship among β, τ, and ρ, we compare $m_\beta(t)$, $m_\tau(8t/9)$, and $m_\rho(4t/3)$ for $|t|$ in [0, 1/4] in Table 2.

As with normal copulas, the medial correlation coefficient is substantially better (for $|t|$ in [0, 1/4]) than either Kendall's tau or Spearman's rho in narrowing the bounds. A similar relationship among β, τ, and ρ holds, at least approximately for values of θ near 0, in families of copulas for which FGM copulas are first-order approximations. Examples of such families include the Ali–Mikhail–Haq family, the Frank family, and the Plackett family. See Nelsen (1999) for details.

4. CONCLUSION

There is a simple geometric explanation for the effectiveness of the medial correlation coefficient in narrowing the Frechet–Hoeffding bounds. Recall that all six of the bounds, $\underline{B}_t$, $\overline{B}_t$, $\underline{T}_t$, $\overline{T}_t$, $\underline{P}_t$, and P_t, are copulas, and hence, as a result of Nelsen et al. (2001), coincide on the boundary of the unit square $\mathbf{I}^2$; and as a result of Nelsen (1999), are uniformly continuous. However, only $\underline{B}_t$ and $\overline{B}_t$ also coincide at the center of the square, i.e., $\underline{B}_t(1/2, 1/2) = \overline{B}_t(1/2, 1/2) = (t+1)/4$. Because m_β, m_τ, and m_ρ are each based on the volume between the graphs of the lower and upper bounds, it is reasonable to expect a smaller volume between the graphs of $\underline{B}_t$ and $\overline{B}_t$ than between $\underline{T}_t$ and $\overline{T}_t$, or between $\underline{P}_t$ and $\overline{P}_t$.

REFERENCES

Blomqvist, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Statist.* 21:593–600.

Kruskal, W. H. (1958). Ordinal measures of association. *J. Am. Statist. Assoc.* 53:814–861.

Nelsen, R. B. (1999). *An Introduction to Copulas*. New York: Springer.

Nelsen, R. B. Quesada Molina, J. J., Rodríguez Lallena, J. A. Úbeda Flores, M. (2001). Bounds on bivariate distribution functions with given margins and measures of association. *Commun. Statist.—Theory Methods* 30:1155–1162.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Request Permission or Order Reprints Instantly!

Interested in copying and sharing this article? In most cases, U.S. Copyright Law requires that you get permission from the article's rightsholder before using copyrighted content.

All information and materials found in this article, including but not limited to text, trademarks, patents, logos, graphics and images (the "Materials"), are the copyrighted works and other forms of intellectual property of Marcel Dekker, Inc., or its licensors. All rights not expressly granted are reserved.

Get permission to lawfully reproduce and distribute the Materials or order reprints quickly and painlessly. Simply click on the "Request Permission/ Order Reprints" link below and follow the instructions. Visit the U.S. Copyright Office for information on Fair Use limitations of U.S. copyright law. Please refer to The Association of American Publishers' (AAP) website for guidelines on Fair Use in the Classroom.

The Materials are for your personal use only and cannot be reformatted, reposted, resold or distributed by electronic means or otherwise without permission from Marcel Dekker, Inc. Marcel Dekker, Inc. grants you the limited right to display the Materials only on your personal computer or personal wireless device, and to copy and download single copies of such Materials provided that any copyright, trademark or other notice appearing on such Materials is also retained by, displayed, copied or downloaded as part of the Materials and is not removed or obscured, and provided you do not edit, modify, alter or enhance the Materials. Please refer to our Website User Agreement for more details.

Request Permission/Order Reprints

Reprints of this article can also be ordered at
http://www.dekker.com/servlet/product/DOI/101081STA200031367