A stochastic ordering based on a decomposition of Kendall's tau

P. CAPÉRAÀ, A.-L. FOUGÈRES & C. GENEST
Département de mathématiques et de statistique
Université Laval
Québec, Canada G1K 7P4

Abstract

If X and Y are random variables with joint distribution function H, their dependence may be measured by Kendall's tau, $\tau(X,Y)$, expressed as $4E(V)-1$ in terms of the random quantity $V=H(X,Y)$ with distribution K on the interval $[0,1]$. A new dependence ordering based on K is defined and studied; although it does not always imply the classical positive quadrant dependence ordering, it is shown to be weaker than the association ordering of Schriever (1987) under weak regularity conditions.

1 Introduction

Given a random variable Z with continuous distribution function H on $\mathbb{R}$, it is well known that $V=H(Z)$ is uniformly distributed on $[0,1]$. This property of the probability integral transformation does not extend to higher dimensions, however. Suppose for example that $Z=(X,Y)$ has continuous bivariate distribution with marginals F and G. If X and Y are mutually independent, then

$$P(V \le v) = P\{F(X)G(Y) \le v\} = K_{\perp}(v) = v - v\log(v), 0 \le v \le 1,$$

while if Y is an almost surely increasing function of X, one has

$$P(V \le v) = P\left[\min\{F(X), G(Y)\} \le v\right] = K_{\min}(v) = v, 0 \le v \le 1.$$

Similarly, one finds

$$P(V \leq v) = P\left[\max\{0, F(X) + G(Y) - 1\} \leq v\right] = K_{\max}(v) = 1, 0 \leq v \leq 1,$$

when Y is an almost surely decreasing function of X.

Assuming that the marginals of X and Y are continuous, let C denote the unique copula associated to H through the relation $H(x, y) = C\{F(x), G(y)\}$. The distribution function K of $V = H(X, Y)$ is clearly a one-dimensional projection of C, and as such, it reflects part of the dependence structure of H that is invariant under strictly increasing transformations of X and Y. Furthermore, Genest & Rivest (1993) point out that K may be regarded as a decomposition of Kendall's coefficient of concordance, τ, since one has

$$\int_0^1 t dK(t) = \int_0^1 \{1 - K(t)\} dt = (\tau + 1)/4. \tag{1}$$

Finally, K sometimes characterizes completely the dependence structure of a copula, such as when it is Archimedean for example (Genest & Rivest, 1993).

In view of the attractive properties of the generalized probability integral transformation $V = H(X, Y)$, it may be tempting to define a dependence ordering for random pairs (X, Y) based on the distribution K of $H(X, Y)$. The purpose of this note is to explore this idea briefly. The exact definition of the new ordering is given and illustrated in §2. It is then shown in §3 that this weak order does not always imply the classical positive quadrant dependence ordering, but that it is weaker than the association ordering of Schriever (1987). A short discussion concludes the paper.

2 Definition of the proposed ordering

For $i = 1, 2$, let (X_i, Y_i) be an observation from copula C_i, and denote by K_i the distribution function of $V_i = C_i(X_i, Y_i)$. Set

$$(X_1, Y_1) \prec (X_2, Y_2) \Leftrightarrow \quad \forall 0 \leq v \leq 1 \quad K_1(v) \geq K_2(v).$$

It is clear that the relation $\prec$ is reflexive, transitive and admits the Fréchet bounds as extrema, since

$$\forall 0 \leq v \leq 1 \quad K_{\min}(v) \leq K(v) \leq K_{\max}(v).$$

In view of equation (1), one also has

$$(X_1, Y_1) \prec (X_2, Y_2) \Rightarrow \tau(X_1, Y_1) \leq \tau(X_2, Y_2),$$

so that $\prec$ may be viewed as a dependence ordering.

Example 1. Denote by C_ϕ the Archimedean copula $\phi^{-1}\{\phi(x) + \phi(y)\}$ generated by a convex, decreasing function $\phi : (0, 1] \rightarrow [0, \infty)$ with the property that $\phi(1) = 0$. Genest & Rivest (1993) show that

$$K_\phi(v) = v - \phi(v)/\phi'(v), 0 < v \leq 1.$$

Consequently, $C_{\phi_1} \prec C_{\phi_2}$ occurs if and only if ϕ_1/ϕ_2 is increasing. As this is the sufficient condition suggested by Genest & MacKay (1986) to check that C_{ϕ_1} is more positive quadrant dependent than C_{ϕ_2}, it is immediate that $\prec$ is stronger than this classical dependence ordering for Archimedean copulas.

Example 2. According to Pickands (1981), bivariate extreme value copulas may be written as

$$C_A(x, y) = \exp\left[\log(xy) A\left\{\log(x)/\log(xy)\right\}\right], 0 \leq x, y \leq 1$$

in terms of a convex function $A : [0, 1] \rightarrow [1/2, 1]$ with the property that $\max(t, 1-t) \leq A(t)$ everywhere. In his thesis, Khoudraji (1995) shows that $K_A(v) = v + (\tau - 1)v \log(v)$, where

$$\tau = \int_0^1 \frac{t(1-t)}{A(t)} dA'(t). \tag{2}$$

Consequently, $C_{A_1} \prec C_{A_2}$ occurs if and only if $\tau_{A_1} \leq \tau_{A_2}$, which implies that positive quadrant dependence is stronger than $\prec$ for extreme value copulas.

The connection between $\prec$ and other dependence orderings will be further studied in the following section.

3 Connection between $\prec$ and other dependence orderings

It follows immediately from formula (2) that for extreme value copulas, $\tau_{A_1} = \tau_{A_2}$ whenever $A_2(t) = A_1(1-t)$ on $[0, 1]$. Using this fact, it is easy to check that one may have $C_{A_1} \prec C_{A_2}$ for extreme value copulas that are not ordered by positive quadrant dependence. It is also likely that examples of copulas can be found that are ordered by positive quadrant dependence, but not by $\prec$. At this point, however, the authors have not been able to identify one explicitly.

The following proposition implies that although the two orderings are not comparable, they are both implied by the associated ordering $\prec_a$ of Schriever (1987), whose definition is recalled below for convenience.

Definition. *The pair (X_2, Y_2) is said to be more associated than (X_1, Y_1), denoted by $(X_1, Y_1) \prec_a (X_2, Y_2)$ or $C_1 \prec_a C_2$, if there exist functions $\psi_i : [0,1]^2 \to [0,1]$, $i = 1, 2$ that are nondecreasing in both arguments, such that (X_2, Y_2) has the same distribution as $(\psi_1(X_1, Y_1), \psi_2(X_1, Y_1))$ and*

$$(-1)^{i-1}\psi_i(x_1, y_1) < (-1)^{i-1}\psi_i(x_2, y_2), i = 1, 2 \Rightarrow x_1 < x_2 \text{ and } y_1 > y_2.$$

Proposition. *For $i = 1, 2$, suppose that (a) copula C_i has a continuous density on $[0,1]^2$ which is strictly positive in the interior of that set; (b) the univariate projection K_i of C_i has density k_i verifying*

$$k_i(t) = o\left\{t^{-1/2} \log^{-1/2-\epsilon_i}(1/t)\right\}$$

for some $\epsilon_i > 0$ as $t \to 0$. Then $C_1 \prec_a C_2 \Rightarrow C_1 \prec C_2$.

The conditions under which this proposition is stated are slightly stronger than those given by Barbe et al. (1996) to guarantee the convergence of the empirical version of K_i, defined in terms of a random sample $(X_{i,1}, Y_{i,1}), \ldots,$ $(X_{i,n}, Y_{i,n})$ from distribution C_i as

$$K_{i,n}(v) = \frac{1}{n} \sum_{j=1}^{n} 1_{\mathbb{R}_+} \{v - C_{i,n}(X_{i,j}, Y_{i,j})\}, 0 \le v \le 1$$

where $C_{i,n}$ is the empirical counterpart of C_i, $i = 1, 2$. Now it is known from the work of Schriever (1987) that

$$C_1 \prec_a C_2 \Rightarrow \forall n \ge 1 \quad C_{1,n} \prec_a C^*_{2,n},$$

where $C^*_{2,n}$ is the empirical distribution function corresponding to the sample $(\psi_1(X_{1,j}, Y_{1,j}), \psi_2(X_{1,j}, Y_{1,j}))$, $1 \le j \le n$. Since $C^*_{2,n}$ has the same distribution as $C_{2,n}$, the proposition will follow by a standard limiting argument if it can be shown that

$$C_{1,n} \prec_a C^*_{2,n} \Rightarrow \forall 0 \le v \le 1 \quad K_{1,n}(v) \ge K^*_{2,n}(v) \tag{3}$$

for arbitrary integer $n \ge 1$.

The following easily verified observation will be useful in establishing implication (3).

Observation. *Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a random sample from copula C with $X_1 < \cdots < X_n$, denote by $\ell = (\ell_1, \ldots, \ell_n)$ the ranks of the Y_i's, and let*

$$N_\ell(k) = \#\left\{1 \le i \le n : \#\{1 \le j < i : \ell_j < \ell_i\} = k-1\right\},$$

for all integer $0 \le k \le n$. Then for all $0 \le k \le n$, one has

$$K_\ell\left(k/n\right) = \frac{1}{n}\sum_{j=1}^{n} 1_{I\!R_+}\left\{\frac{k}{n} - C_n(X_i, Y_i)\right\} = \frac{N_\ell(0) + \cdots + N_\ell(k)}{n}. \tag{4}$$

To prove (3), take $C_{1,n} \prec_a C^*_{2,n}$ and suppose without loss of generality that the random samples from which they were constructed both have their X values ordered. Then according to Theorem 3.2 of Block et al. (1990), the vector ℓ_2 of ranks of the Y_j's from the second sample may be obtained from the vector ℓ_1 of ranks of the Y_i's of the first sample in a finite number of steps, each of which consists of correcting an inversion of type 2 or 3, as defined by these authors. Thus it suffices to check that *each* such correction causes a drop in $K_{1,n}(v)$ over its entire domain.

To see why this is true for a correction of type 3, for example, assume that $\ell_{1,k} < \ell_{1,k-1}$ for some $1 < k \le n$. By definition, the correction then consists of permuting these two values, which leads to a new vector of ranks ℓ^*_1 differing from ℓ_1 in that $\ell^*_{1,k-1} = \ell_{1,k}$ and $\ell^*_{1,k} = \ell_{1,k-1}$. Setting

$$\alpha = \#\{1 \le i < k-1 : \ell_{1,i} < \ell_{1,k-1}\} \text{ and } \beta = \#\{1 \le i < k : \ell_{1,i} < \ell_{1,k}\},$$

one has either $\alpha = \beta$ or $\alpha > \beta$ by hypothesis. An elementary counting argument then shows that in both cases, $N_{\ell^*}(k) = N_\ell(k)$ for all $0 \le k \le n$, except

$$N_{\ell^*}(\alpha) = N_\ell(\alpha) - 1 \text{ and } N_{\ell^*}(\alpha+1) = N_\ell(\alpha+1) + 1.$$

Thus it follows from formula (4) that $K_{1,n}(v) = K_{\ell_1}(v) \ge K_{\ell^*}(v)$ for all $0 \le v \le 1$, as required.

4 Conclusion

Although the above discussion was restricted to the bivariate case, it is clear that the ordering $\prec$ makes sense in higher dimensions as well. In particular, it is easy to see that this extension orders the generalized version of Kendall's tau defined by Joe (1990). An obvious question for future research is whether the regularity conditions under which the proposition was established can be eliminated. Another avenue to explore would be to define an ordering based on some decomposition of Spearman's rho.

Acknowledgements

Funds in partial support of this work were supplied by the Natural Sciences and Engineering Research Council of Canada and the Fonds pour la formation de chercheurs et l'aide à la recherche du Gouvernement du Québec.

References

Barbe, P., C. Genest, K. Ghoudi and B. Rémillard (1996). On Kendall's process. *Journal of Multivariate Analysis*, **58**, 197-229.

Block, H. W., D. Chhetry, Z. Fang and A. R. Sampson (1990). Partial orders on permutations and dependence orderings on bivariate empirical distributions. *The Annals of Statistics*, **18**, 1840-1850.

Genest, C. and R. J. MacKay (1986). Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *The Canadian Journal of Statistics*, **14**, 145-159.

Genest, C. and L.-P. Rivest (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*, **88**, 1034-1043.

Joe, H. (1990). Multivariate concordance. *Journal of Multivariate Analysis*, **35**, 12-30.

Khoudraji, A. (1995). *Contributions à l'étude des copules et à la modélisation de valeurs extrêmes bivariées.* Unpublished doctoral dissertation, Université Laval, Québec, Canada.

Pickands, J. (1981). Multivariate extreme value distributions. *Bulletin of the International Statistical Institute*, 859-78.

Schriever, B. F. (1987). An ordering of positive dependence. *The Annals of Statistics*, **15**, 1208-1214.