NOTES

Copulas, Characterization, Correlation, and Counterexamples

ROGER B. NELSEN
Lewis and Clark College
Portland, OR 97219-7899

1. Copulas Copulas are functions that join univariate distribution functions to form multivariate distribution functions. They were first introduced in 1959 by A. Sklar **[10]** to answer some questions posed by M. Fréchet concerning the relationship between a multidimensional probability distribution function and its lower dimensional marginals. Over the past 30 years or so, copulas have played an important role in several areas of probability and statistics, including multivariate distribution theory, nonparametric statistics, and Markov processes; but only recently have they come to the attention of the general statistical and mathematical communities. In this paper, we will survey some of the important properties of copulas, and demonstrate ways in which copulas can be employed in probability and mathematical statistics courses to enhance and illuminate the presentation of a number of topics. Specifically, we will show how copulas can be used in statistics i) to characterize some dependence concepts for two random variables; ii) to obtain a geometric interpretation of the population version of a nonparametric correlation coefficient, and to illustrate how that coefficient measures dependence; and iii) to facilitate the generation of counterexamples.

A (two-dimensional) *copula* is a function $C: \mathbf{I}^2 \to \mathbf{I} = [0,1]$ with the following properties:

(i) $C(0,t) = C(t,0) = 0$ and $C(1,t) = C(t,1) = t$ for all t in $\mathbf{I}$; and (1.1)
(ii) $C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_1) + C(u_1,v_1) \geq 0$ for all u_1, u_2, v_1, v_2 in $\mathbf{I}$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$.

It follows that C is nondecreasing in each variable (let $v_1 = 0$ or $u_1 = 0$ above) and continuous (since (1.1) implies that C satisfies the Lipschitz condition $|C(u_2,v_2) - C(u_1,v_1)| \leq |u_2 - u_1| + |v_2 - v_1|$). If one thinks of $C(u,v)$ as assigning a number in $\mathbf{I}$ to the rectangle $[0,u] \times [0,v]$ then part (ii) of (1.1) gives an "inclusion-exclusion"-type formula for the number assigned by C to each rectangle $[u_1,u_2] \times [v_1,v_2]$ in $\mathbf{I}^2$, and states that the number so assigned must be nonnegative.

The importance of copulas to mathematical statistics is described in

Sklar's Theorem. *Let H be a two-dimensional distribution function (d.f.) with marginal d.f.'s F and G. Then there exists a copula C such that $H(x,y) = C(F(x),G(y))$. Conversely, for any univariate d.f.'s F and G and any copula C, the function H defined above is a two-dimensional d.f. with marginals F and G. Furthermore, if F and G are continuous, C is unique.*

Thus a copula "couples" a bivariate d.f. to its one-dimensional marginal d.f.'s. A proof of Sklar's theorem may be found in [8]. In addition, a copula is itself a bivariate distribution function with marginals uniform on $\mathbf{I}$.

As noted by M. Fréchet [1], it is an elementary exercise to show that any bivariate d.f. H with marginal d.f.'s F and G satisfies

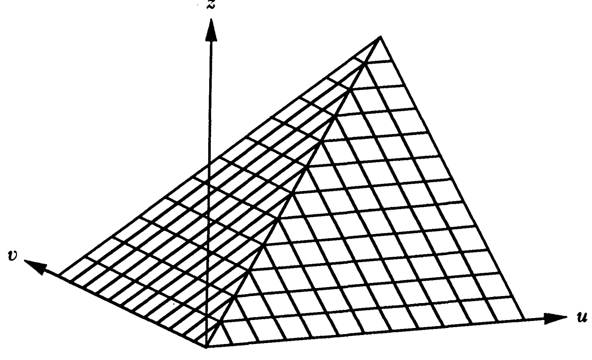
$$\max\{F(x)+G(y)-1,0\} \le H(x,y) \le \min\{F(x),G(y)\}. \tag{1.2}$$

To see this, observe that: $H(x,y)=\Pr\{X\le x, Y\le y\}\le \Pr\{X\le x\}=F(x)$; similarly $H(x,y)\le G(y)$; and $1\ge \Pr\{X\le x \text{ or } Y\le y\}=F(x)+G(y)-H(x,y)$. The bounds in (1.2) for bivariate d.f.'s are often referred to as the lower and upper *Fréchet bounds*, *and are themselves bivariate d.f.'s. In terms of copulas, (1.2) is*

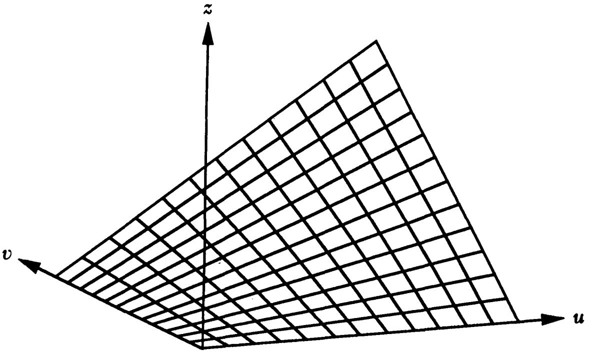
$$\max\{u+v-1,0\} \le C(u,v) \le \min\{u,v\} \quad \text{for all } u,v\in \mathbf{I}, \tag{1.3}$$

and the lower and upper bounds in (1.3) are the copulas for the Fréchet bounds in (1.2).

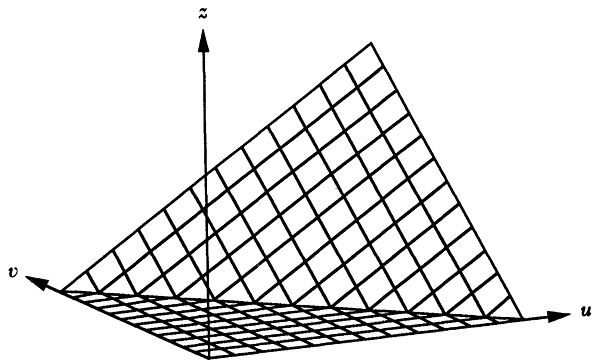
Several of our applications involve the shape of the graph of a copula, i.e., of the surface $z=C(u,v)$. It follows from (1.1) that the graph of any copula is a continuous surface whose boundary is the skew quadrilateral with vertices $(0,0,0)$, $(0,1,0)$, $(1,1,1)$, and $(1,0,0)$; and from (1.3) that this graph lies between the surfaces $z=\max\{u+v-1,0\}$ and $z=\min\{u,v\}$. In the figures we display the graphs of the copulas $\max\{u+v-1,0\}$ and $\min\{u,v\}$ for the Fréchet bounds, and the important intermediate case, $z=C(u,v)=uv$. Alternatively, the graph of a copula can be viewed as the graph of the joint d.f. $z=H(x,y)$ in which the x and y axes have been relabelled in units of $u=F(x)$ and $v=G(y)$. For a complete historical discussion of copulas and an extensive bibliography, see [7].



$z=\min\{u,v\}$



$z=uv$



$z = \max\{u + v - 1, 0\}$

2. Characterization In this section and the next, let X and Y be continuous random variables (r.v.'s) with joint d.f. H, marginal d.f.'s, F and G, respectively, and copula C. Many properties of the pair X, Y may be succinctly expressed as properties of the corresponding copula C. Independence is one such property: in terms of d.f.'s X and Y are independent if, and only if, $H(x, y) = F(x)G(y)$. So, as a consequence of Sklar's theorem, every pair of independent r.v.'s has the same copula, or

(a) X and Y are independent if, and only if, $C(u, v) = uv$.

At the other extreme, monotone functional dependence of r.v.'s can also be shown to be a property of the copula: Specifically,

(b) Y is almost surely an increasing function of X if, and only if, $C(u, v) = \min\{u, v\}$; and

(c) Y is almost surely a decreasing function of X if, and only if, $C(u, v) = \max\{u + v - 1, 0\}$.

These last two results are usually attributed to Fréchet. For a recent proof, see [3].

Another statistical concept similar to but weaker than independence is *exchangeability*. A (finite) set of r.v.'s is exchangeable if the individual r.v.'s are identically distributed; every pair of r.v.'s has the same joint distribution as every other pair; and so on. Whereas independent r.v.'s are exchangeable, exchangeable r.v.'s need not be independent. For example, sampling with replacement from a finite population yields observations that are independent and exchangeable, sampling without replacement yields observations that are exchangeable but not independent. In the bivariate setting, X and Y are exchangeable if the vectors (X, Y) and (Y, X) have the same distribution. So exchangeability of r.v.'s is essentially equivalent to symmetry of their copula, or:

(d) X and Y are exchangeable if, and only if, $F = G$ and $C(u, v) = C(v, u)$.

In a number of statistical situations, it may be desirable to describe qualitatively the statement that "large values of Y occur with large values of X, and small values of Y occur with small values of X." For example, if X and Y denote a student's percentile scores on two achievement tests, then X and Y may tend to be simultaneously high or simultaneously low. One such "positive dependence property" is *positive quadrant dependence* [2]: X and Y are positively quadrant dependent if their joint distribution is such that the probability that they are simultaneously "small" is at least as great as it would be were they independent r.v.'s; that is, $\Pr\{X \leq x, Y \leq y\} \geq \Pr\{X \leq x\}\Pr\{Y \leq y\}$ for all x and y. Since this is equivalent to $H(x, y) \geq F(x)G(y)$, it readily follows that

(e) X and Y are positively quadrant dependent if, and only if, $C(u, v) \geq uv$ for all $u, v \in \mathbf{I}$. A geometric interpretation of positive quadrant dependence is immediate: X

and Y are positively quadrant dependent if, and only if, the graph of $z = C(u,v)$ lies on or above the graph of $z = uv$.

For a discussion of several other dependence concepts in terms of copulas and their corresponding geometric interpretations, see [4].

3. Correlation As we've seen in the preceding section, a property of r.v.'s that doesn't depend on the form of the marginal distributions can often be expressed in terms of the copula. Indeed, it can be shown [9] that if f and g are strictly increasing (almost surely) on the ranges of X and Y, respectively, then the copula of the r.v.'s $f(X)$ and $g(Y)$ is the same as the copula of X and Y. In other words, while almost surely strictly increasing functions of the r.v.'s will alter the marginal and joint d.f.'s, they leave the copula unchanged. Thus it is the copula that expresses the "nonparametric" or "distribution-free" properties of the joint distribution of X and Y, i.e., those properties (such as positive, quadrant dependence) that are invariant under almost surely strictly increasing transformations. Hence one might suspect that the population analog of a nonparametric correlation coefficient such as Spearman's rho (ρ_s) would be expressible in terms of the copula. Such is indeed the case, and the result yields a nice geometric interpretation of ρ_s.

Spearman's ρ_s, introduced by the psychologist C. Spearman in 1904, is also known as the grade correlation coefficient, a term introduced by K. Pearson. The population version of Spearman's ρ_s can be defined as the ordinary Pearson product-moment correlation coefficient, applied not to the r.v.'s X and Y, but to their "grades" $U = F(X)$ and $V = G(Y)$. Since U and V are uniform on $\mathbf{I}$ (note that $F(x) = \Pr\{X \le x\} = \Pr\{U \le F(x)\}$), we have $E(U) = E(V) = 1/2$ and $\operatorname{Var}(U) = \operatorname{Var}(V) = 1/12$. The joint d.f. of U and V is C, and thus

$$\rho_s = \frac{E(UV) - 1/4}{1/12} = 12E(UV) - 3 = 12\iint_{\mathbf{I}^2} uv\, dC - 3.$$

But if (U', V') is any other pair of r.v.'s independent of (U,V), each uniform on $(0,1)$ with joint d.f. C', then $\iint_{\mathbf{I}^2} C(u,v)\, dC' = \iint_{\mathbf{I}^2} C'(u,v)\, dC$ since

$$\Pr\{U \le U', V \le V'\} = \iint_{\mathbf{I}^2} \Pr\{U \le u, V \le v\}\, dC' = \iint_{\mathbf{I}^2} C(u,v)\, dC',$$

and

$$\begin{aligned}\Pr\{U' \ge U, V' \ge V\} &= \iint_{\mathbf{I}^2} \Pr\{U' \ge u, V' \ge v\}\, dC \\ &= \iint_{\mathbf{I}^2} [1 - u - v + C'(u,v)]\, dC = \iint_{\mathbf{I}^2} C'(u,v)\, dC.\end{aligned}$$

Thus $\iint_{\mathbf{I}^2} uv\, dC = \iint_{\mathbf{I}^2} C\, dudv$ and hence

$$\rho_s = 12\iint_{\mathbf{I}^2} C\, dudv - 3 = 12\iint_{\mathbf{I}^2} [C(u,v) - uv]\, dudv. \tag{3.1}$$

From (3.1) we obtain simple geometric interpretations of Spearman's ρ_s. Since $\iint_{\mathbf{I}^2} C\, dudv$ represents the volume of the portion of the unit cube below the graph of $z = C(u,v)$, we see from (1.3) that the value of this integral lies between $1/6$ and $1/3$. Thus ρ_s can be viewed as the volume under the graph of $z = C(u,v)$ over $\mathbf{I}^2$, scaled to lie between -1 and $+1$, or as the (scaled) signed volume between the surfaces $z = C(u,v)$ and $z = uv$. Indeed, any measure of distance between these surfaces is a measure of dependence. See [**9**] for further examples.

In the previous section we observed that X and Y are positively quadrant dependent if, and only if, $C(u,v) \geq uv$ on $\mathbf{I}^2$. So, in a sense, the quantity $C(u,v) - uv$ is a measure of "local" positive (and negative) quadrant dependence, and thus $\rho_s/12 = \iint_{\mathbf{I}^2}[C(u,v) - uv]\,dudv$ can be interpreted as a measure of "average" quadrant dependence.

Similar results exist for the population version of another nonparametric correlation coefficient, Kendall's tau. For details, see [**5**].

4. Counterexamples When the student of probability and statistics encounters a statement, the task usually is (as in other branches of mathematics) to either provide a proof if it is true or find a counterexample if it is false. In this final section we will illustrate the ease with which copulas can be used to construct counterexamples. Of course, other examples for each of the following situations exist; see, for example, [**6**] and [**11**].

For the first five examples, consider the copula formed by averaging the two copulas for the Fréchet bounds, i.e., $C_1(u,v) = (1/2)(\max\{u+v-1,0\} + \min\{u,v\})$. Let U and V be r.v.'s with joint d.f. C_1. This copula concentrates the probability mass uniformly on the two diagonals of $\mathbf{I}^2$, i.e., $(U-V)(U+V-1) = 0$ holds almost surely. Whenever there is at least one relationship $\varphi(U,V) = 0$ that holds with probability 1 on $\mathbf{I}^2$ (but φ not identically 0 on $\mathbf{I}^2$), we call the distribution (and the copula) *singular*. We note that the Fréchet bounds are singular as well.

Example 1. There exist exchangeable random variables that are not independent. Note that U and V have the same marginal distribution and that $C_1(u,v) = C_1(v,u)$. But more is true—U and V are also uncorrelated, since $\operatorname{Cov}(U,V) = E(UV) - E(U)E(V) = 1/4 - (1/2)^2 = 0$.

Example 2. There is a bivariate distribution without a density but whose marginals possess densities (note that $\partial^2 C_1(u,v)/\partial u \partial v = 0$ almost everywhere).

For the next three examples, let X and Y be standard normal r.v.'s with d.f. Φ.

Example 3. There exist bivariate distributions with normal marginals that are not bivariate normal. Indeed, if $N_\rho(x,y)$ denotes a standard bivariate normal d.f. with (Pearson) correlation coefficient ρ for some $\rho \in (-1,1)$, then *any* copula *except* one of the form $C(u,v) = N_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$ will suffice!

Example 4. There exist uncorrelated normal random variables that are not independent (note that $E(XY) = 0$ but $C_1(u,v) \neq uv$). Recall that if X and Y are bivariate normal, then X and Y are independent if, and only if, they are uncorrelated —and this example shows that this conclusion need not hold when the joint d.f. is not a bivariate normal.

Example 5. Two normal random variables may have a sum that is not normal (note that $\Pr\{X+Y=0\} = 1/2$). Again, if X and Y are bivariate normal, then the sum (or any other linear combination) is also normal.

For the final four examples, let U and V be r.v.'s uniform on $\mathbf{I}$, but satisfying $V = |2U-1|$, and let C_2 denote their copula. To find an expression for C_2, we evaluate $C_2(u,v) = \Pr\{U \leq u, V \leq v\}$ where $V = 1-2U$ when $u \in [0,1/2]$ and $V = 2U-1$ when $u \in (1/2,1]$. This yields

$$C_2(u,v) = \begin{cases} \max\left\{u + \frac{1}{2}(v-1), 0\right\}, u \in \left[0, \frac{1}{2}\right], \\ \min\left\{u + \frac{1}{2}(v-1), v\right\}, \ u \in \left(\frac{1}{2}, 1\right]. \end{cases}$$

This copula, again singular, concentrates the probability uniformly on the polygonal line in $\mathbf{I}^2$ that joins $(0,1)$ to $(1/2,0)$ to $(1,1)$, i.e., the graph of $v=|2u-1|$ for $u \in \mathbf{I}$.

Example 6. Two r.v.'s can be uncorrelated although one can be predicted perfectly from the other ($\mathrm{Cov}(U,V)=0$ but $V=|2U-1|$).

Example 7. Two r.v.'s may be identically distributed and uncorrelated but not exchangeable (C_2 is not symmetric in u and v).

Example 8. There exist identically distributed r.v.'s whose difference is not symmetric about 0 (note that $\Pr\{U-V>0\}=\Pr\{U>1/3\}=2/3$).

Example 9. There exist pairs of r.v.'s each symmetric about 0, but whose sum is not symmetric about 0 (Let $X=2U-1$ and $Y=2V-1$ so that X and Y are uniform on $(-1,1)$, then note that $\Pr\{X+Y>0\}=\Pr\{U+V>1\}=\Pr\{U>2/3\}=1/3$).

REFERENCES

1. M. Fréchet, Sur les tableaux de corrélation dont les marges sont données, *Annales de l'Université de Lyon*, 9 (1951), Sect. A, 53–77.
2. E. L. Lehmann, Some concepts of dependence, *Annals of Mathematical Statistics*, 37 (1966), 1137–1153.
3. P. Mikusiński, H. Sherwood, and M. D. Taylor, The Fréchet bounds revisited, *Real Analysis Exchange*, 17 (1991–92), 759–764.
4. R. B. Nelsen, Copulas and association, in *Advances in Probability Distributions with Given Marginals*, eds. G. Dall'Aglio, S. Kotz and G. Salinetti; Kluwer Academic Publishers, Dordrecht (1991), 51–74.
5. R. B. Nelsen, On measures of association as measures of positive dependence, *Statistics & Probability Letters* 14 (1992), 269–274.
6. J. P. Romano and A. F. Siegel, *Counterexamples in Probability and Statistics*, Wadsworth and Brooks/ Cole, Pacific Grove, CA (1986).
7. B. Schweizer, Thirty years of copulas, in *Advances in Probability Distributions with Given Marginals*, eds. G. Dall'Aglio, S. Kotz and G. Salinetti; Kluwer Academic Publishers Dordrecht (1991), 13–50.
8. B. Schweizer and A. Sklar, *Probabilistic Metric Spaces*, Elsevier Science Publishing Co., New York, 1983.
9. B. Schweizer and E. F. Wolff, On nonparametric measures of dependence for random variables, *Annals of Statistics*, 9 (1981), 879–885.
10. A. Sklar, Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut Statistique de l' Université de Paris*, 8 (1959), 229–231.
11. J. M. Stoyanov, *Counterexamples in Probability*, John Wiley & Sons, Inc., New York, 1987.