Stochastica XIII-1 (1992), 61-74

SHUFFLES OF MIN

P. MIKUSINSKI, H. SHERWOOD & M.D. TAYLOR

ABSTRACT

Copulas are functions which join the margins to produce a joint distribution function. A special class of copulas called shuffles of Min is shown to be dense in the collection of all copulas. Each shuffle of Min is interpreted probabilistically. Using the above-mentioned results, it is proved that the joint distribution of any two continuously distributed random variables X and Y can be approximated uniformly, arbitrarily closely by the joint distribution of another pair X^* and Y^* each of which is almost surely an invertible function of the other such that X and X^* are identically distributed as are Y and Y^*. The preceding results shed light on A. Rényi's axioms for a measure of dependence and a modification of those axioms as given by B. Schweizer and E.F. Wolff.

AMS 1980 Subject Classification (1985 revision). Primary 60E05, Secondary 60B10, 62E10, 62H05, 62H20.

Key Words and Phrases. Copula, doubly stochastic measure, functional dependence, stochastic independence, measure of dependence.

1. Introduction.

Throughout this paper X and Y denote continuously distributed random variables; F, G and H denote the distribution functions of X, Y and (X, Y), respectively. If X and Y bear affixes, then F, G and H bear these same affixes. Also, I, R and $\overline{\mathsf{R}}$ denote the closed unit interal [0,1], the reals and the extended reals, respectively.

In 1959 A. Sklar [8] introduced the notion of a copula which is a function from I^2 into I satisfying

$$C(s,0) = 0 = C(0,s) \quad \text{and} \quad C(s,1) = s = C(1,s) \tag{1.1}$$

whenever $0 \le s \le 1$, and

$$C(s_2,t_2) - C(s_2,t_1) + C(s_1,t_1) - C(s_1,t_2) \ge 0 \tag{1.2}$$

whenever $0 \le s_1 \le s_2 \le 1$ and $0 \le t_1 \le t_2 \le 1$.

Some copulas of particular importance are π, M and W where for any (s,t) in I^2, we have $\pi(s,t) = st$, $M(s,t) = \mathrm{Min}(s,t)$ and $W(s,t) = \mathrm{Max}(s+t-1, 0)$. An extensive treatment of copulas is given by B. Schweizer and Sklar in [6] where they show in their Lemma 6.1.9 that for any s, s', t, t' in I and any copula C,

$$|C(s,t) - C(s',t')| \le |s-s'| + |t-t'|. \tag{1.3}$$

Corresponding to X and Y there is a unique copula C, called the connecting copula for (X,Y), such that

$$H(x,y) = C(F(x), G(y)) \quad \text{for all } x, y \text{ in } \mathbb{R}. \tag{1.4}$$

(If F or G fails to be continuous, the word "unique" must be deleted in the preceding statement).

Sklar's paper was preceded by a paper of M. Fréchet [2] in which Fréchet presents the following result which we reformulate using Sklar's terminology: (X,Y) has M, (respectively, W) as its connecting copula if and only if each of X and Y is almost surely an increasing (respectively, decreasing) Borel-measurable function of X. Of course (X,Y) has π as its connecting copula if and only if X and Y are stochastically independent. These three examples illustrate what we mean by a probabilistic interpretation of a copula C.

A copula is a bivariate distribution on I^2 having uniform margins. In particular π has its mass spread uniformly across the entire square while M has all its mass spread uniformly along the diagonal from (0,0) to (1,1) and W has all its mass spread uniformly along the other diagonal.

Of special interest in this paper are the copulas we call shuffles of Min. The mass distribution for a shuffle of Min can be obtained by (1) placing the mass distribution for M(=Min) on I^2, (2) cutting I^2 vertically into a finite number of strips, (3) shuffling the strips with perhaps some of them flipped around their vertical axes of symmetry, and then (4) reassembling them to form the square again. The resulting mass distribution will correspond to a copula called a <u>shuffle of Min</u>. In section two we probabilistically interpret these shuffles of Min.

A doubly stochastic measure (briefly, a dsm) is a measure μ defined at least on the Borel subsets of I having the property that $\mu(A \times I) = \mu(I \times A) = m(A)$ where A is any Borel subset of I and m denotes Lebesgue measure. Given any dsm μ on I^2, the function C defined on I^2 via

$$C(x,y) = \mu([0,x] \times [0,y]) \tag{1.5}$$

is a copula. Or, given a copula C, one may use (1.5) to define μ on sets of the form $[0,x] \times [0,y]$ and extend the definition of μ to the Borel subsets of I^2 by well-known techniques. The resultant μ will be a dsm which, in this paper, we denote by μ_C. We denote by COP (DSM) the set of all copulas (doubly stochastic measures). With reasonable topologies on COP and DSM, the one-one correspondence provided by (1.5) is a homeomorphism [3].

J.R. Brown [1] proves that DSM and the set of all Markov operators on $L_\infty(I)$ are homeomorphic. Moreover with each invertible Lebesgue- measure preserving map ϕ on I he associates a Markov operator T_ϕ and shows that the set of all these T_ϕ's is dense in the set of all Markov operators on $L_\infty(I)$. Consequently the set of copulas corresponding to the T_ϕ's properly contains the shuffles of Min. As shown in section three, however, the set consisting only of shuffles of Min is also dense in COP.

In the final section we use the results of sections two and three to show that for any two continuously distributed random variables X and Y (including stochastically independent ones), there exist two random variables X^* and Y^*, each almost surely an invertible function of the other, such that $F = F^*$, $G = G^*$ while H and H^* are uniformly, arbitrarily close. Lastly, we use the preceding result to show that certain of the Rényi axioms [5] for a measure of dependence for random variables are inconsistent with a continuity- type condition on the measure of dependence which B. Schweizer and E.F. Wolff [7] showed was satisfied by several natural nonparametric measures of dependence.

2. Probabilistic interpretations.

We begin with

Definition 2.1. A copula C is a shuffle of Min if and only if there is a positive integer n, two partitions $0 = s_0 < s_1 < \ldots < s_n = 1$ and $0 = t_0 < t_1 < \ldots < t_n = 1$ of I, and a permutation σ on $\{1, 2, \ldots, n\}$ such that each $[s_{i-1}, s_i] \times [t_{\sigma(i)-1}, t_{\sigma(i)}]$ is a square in which C deposits a mass of size $s_i - s_{i-1}$ spread uniformly along one of the diagonals. For each $i = 1, 2, \ldots, n$ we let $m(i)$ denote the slope of the diagonal of $[s_{i-1}, s_i] \times [t_{\sigma(i)-1}, t_{\sigma(i)}]$ along which the mass in that square is spread. If $m \equiv 1$ we say C is a straight shuffle of Min. If $m \equiv -1$ we say that C is a flipped shuffle of Min. Also, we say that C is the shuffle of Min generated by $(n, \{s_i\}, \{t_i\}, \sigma, m)$.

Before we give our probabilistic interpretation of shuffles of Min we need some terminology. For subsets A and B of $\overline{\mathbb{R}}^2$, we say A is below (strictly below) B if and only if $y_1 \le y_2$ $(y_1 < y_2)$ whenever $(x_1, y_1) \in A$ and $(x_2, y_2) \in B$. We say A is to the left of (strictly to the left of) B if and only if $x_1 \le x_2$ $(x_1 < x_2)$ whenever

$(x_1, y_1) \in A$ and $(x_2, y_2) \in B$.

The following lemma provides machinery which enables us to see geometrically what happens with the probabilistic interpretation of a shuffle of Min. The proof of each part of the Lemma is straighforward so we omit it.

Lemma 2.1. Suppose C is the connecting copula for (X, Y). Define $\phi : \overline{\mathbf{R}}^2 \to I^2$ via $\phi(x, y) = (F(x), G(y))$. Then,

a) $\phi : \overline{\mathbf{R}}^2 \overset{\text{onto}}{\longrightarrow} I^2$.

b) If $A, B \subset \overline{\mathbf{R}}$, then $\phi(A \times B) = F(A) \times G(B)$.

c) If $A, B \subset I$, then $\phi^{-1}(A \times B) = F^{-1}(A) \times G^{-1}(B)$.

d) If $A, B \subset \overline{\mathbf{R}}^2$ and A is below (to the left of) B, then $\phi(A)$ is below (to the left of) $\phi(B)$.

e) If $A, B \subset I^2$ and A is strictly below (strictly to the left of) B, then $\phi^{-1}(A)$ is strictly below (strictly to the left of) $\phi^{-1}(B)$.

f) If $S = [x_1, x_2] \times [y_1, y_2] \subset \overline{\mathbf{R}}^2$, then $P[(X, Y) \in S] = \mu_C(\phi(S))$.

g) If $S = [s_1, s_2] \times [t_1, t_2] \subset \overline{\mathbf{I}}^2$, then $P[(X, Y) \in \phi^{-1}(S)] = \mu_C(S)$.

Definition 2.2. A function $f : \overline{\mathbf{R}} \to \overline{\mathbf{R}}$ is <u>strongly piecewise monotone</u> if and only if there is a positive integer n, two partitions

$$-\infty = x_0 < x_1 < \ldots < x_n = +\infty \text{ and } -\infty = y_0 < y_1 < \ldots < y_n = +\infty \text{ of } \overline{\mathbf{R}},$$

and a permutation σ on $\{1, 2, \ldots, n\}$ such that, for each $i = 1, 2, \ldots, n$, the function $f|_{(x_{i-1}, x_i]}$ is monotone with $f(x_i) \neq f(x_{i-1}+)$ and $\text{Ran}(f|_{(x_{i-1}, x_i]}) \subset |y_{\sigma(i)-1}, y_{\sigma(i)}]$. For each $i = 1, 2, \ldots, n$ we let

$$m(i) = \begin{cases} 1, & \text{if } f(x_i) > f(_{x_i-1}+), \\ -1, & \text{if } f(x_i) < f(_{x_i-1}+), \end{cases}$$

In this case we say f has components $(n, \{x_i\}, \{y_i\}, \sigma, m)$. If $m \equiv 1$ we say f is strongly piecewise nondecreasing. If $m \equiv -1$ we say that f is strongly piecewise non increasing.

We are now ready to give our probabilistic interpretation of shuffles of Min. In one direction we have

Theorem 2.1. Let C be a shuffle of Min generated by $(n, \{s_i\}, \{t_i\}, \sigma, m)$. Suppose C is the connecting copula for (X, Y). Then $Y = f \circ X$ almost surely for some strongly piecewise monotone function f having components $(n, \{x_i\}, \{y_i\}, \sigma, m)$ where, for $1 = 1, 2, \ldots, n-1$, $x_i = \sup\{x : F(x) = s_i\}$ and $y_i = \sup\{y : G(y) = t_i\}$.

Proof: Using part (b) of Lemma 2.1, we obtain that $\phi([x_{i-1}, x_i] \times [y_{j-1}, y_j]) = [s_{i-1}, s_i] \times [t_{j-1}, t_j]$. Now, from Definition 2.1 and part (f) of Lemma 2.1 we obtain

$$P[(X,Y) \in [x_{i-1}, x_i] \times [y_{j-1}, y_j]] = \begin{cases} 0, & \text{if } j \neq \sigma(i), \\ s_i - s_{i-1}, & \text{if } j = \sigma(i), \end{cases}$$

Now fix an $i \in \{1, 2, \ldots, n\}$. let D_i denote the diagonal in $(s_{i-1}, s_i] \times (t_{\sigma(i)-1}, t_{\sigma(i)}]$ with slope $m(i)$. We want to learn what $\phi^{-1}(D_i)$ looks like. There are two cases to be considered: either $m(i) = 1$ or $m(i) = -1$.

Case 1. Suppose $m(i) = 1$. Then

$$\begin{aligned} D_i &= \{(s_{i-1} + r, t_{\sigma(i)-1} + r) : 0 < r \leq s_i - s_{i-1}\} \\ &= \bigcup_{0 < r \leq s_i - s_{i-1}} \left(\{s_{i-1} + r\} \times \{t_{\sigma(i)-1} + r\}\right). \end{aligned}$$

Thus, according to Lemma 2.1(c), $\phi^{-1}(D_i)$ is a union of sets of the form

$$F^{-1}(\{s_{i-1} + r\}) \times G^{-1}\left(\{t_{\sigma(i)-1} + r\}\right) = [a_r, b_r) \times [c_r, d_r]$$

where for all but countably many r in $(0, s_i - s_{i-1}]$, $a_r = b_r$ and $c_r = d_r$. Whenever $r_1 < r_2$ we conclude, using Lemma 2.1(e), that $[a_{r_1}, b_{r_1}] \times [c_{r_1}, d_{r_1}]$ is strictly below and strictly to the left of $[a_{r_2}, b_{r_2}] \times [c_{r_2}, d_{r_2}]$. See Figure 2.1 for a picture of what $\phi^{-1}(D_i)$ may look like.

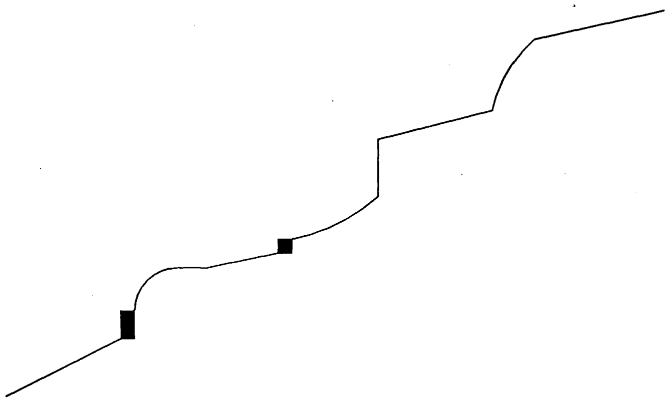

Figure 2.1.

It follows from Lemma 2.1(f), using standard arguments, that $P[(X,Y) \in \phi^{-1}(D_i)] = \mu_C(D_i) = s_i - s_{i-1}$. By deleting from $\phi^{-1}(D_i)$ at most a countable family of sets in which (X,Y) lies with probability zero, we obtain

$$\bigcup_{0<r\leq s_i - s_{i-1}} [a_r, b_r] \times \{c_r\},$$

the graph of a nondecreasing function which we call g_i. It is easy to verify that the domain of g_i is $(x_{i-1}, x_i]$ and the range of g_i is a subset of $[y_{\sigma(i)-1}, y_{\sigma(i)}]$. Since $P[Y = g_i \circ X] = s_i - s_{i-1} > 0$ and G is continuous, it follows that $g_i(x_i) > g_i(x_{i-1}+)$.

Case 2. Suppose $m(i) = -1$. Following a procedure similar to that used in the preceding case we construct a nonincreasing function $g_i : (x_{i-1}, x_i] \to [y_{\sigma(i)-1}, y_{\sigma(i)}]$ such that $P[Y = g_i \circ X] = s_i - s_{i-1}$.

Finally, we define $f : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ as follows: for any $x \in \overline{\mathbb{R}}$ with $-\infty < x$ there is exactly one $i \in \{1, 2, \ldots, n\}$ such that $x_{i-1} < x \leq x_i$; we let $f(x) = g_i(x)$. We define $f(-\infty) = \lim_{x\to-\infty} f(x)$. Clearly f is a strongly piecewise monotone function having components $(n, \{x_i\}, \{y_i\}, \sigma, m)$. To see that $Y = f \circ X$ almost surely observe that

$$P[Y = f \circ X] = \sum_{i=1}^{n} P[Y = g_i \circ X] = \sum_{i=1}^{n} (s_i - s_{i-1}) = 1.$$

This completes the proof.

Theorem 2.2. Suppose $f : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ is a strongly piecewise monotone function having components $(n, \{x_i\}, \{y_i\}, \sigma, m)$. Suppose $Y = f \circ X$ almost surely. Suppose further that for each $i = 1, 2, \ldots, n$, $F(x_i) > F(x_{i-1})$. Then the connecting copula, C, for (X, Y) is the shuffle of Min generated by $(n, \{F(x_i)\}, \{G(y_i)\}, \sigma, m)$.

Proof. From Lemma 2.1(b) it follows that ϕ carries the lines $x = x_i$ and $y = y_j$ in $\mathbb{R}^2$ into the lines $s = F(x_i)$ and $t = G(y_j)$ in I^2, respectively. Moreover, the order of the lines from left to right and from bottom to top is preserved. Parts (b) and (f) of Lemma 2.1 guarantee that

$$\begin{aligned}\mu_C([F(x_{i-1}), F(x_i)] \times [G(y_{\sigma(i)-1}), G(y_{\sigma(i)})]) &= P[(X, Y) \in [x_{i-1}, x_i] \times [y_{\sigma(i)-1}, y_{\sigma(i)}]] \\ &= F(x_i) - F(x_{i-1}) = G(y_{\sigma(i)}) - G(y_{\sigma(i)-1}).\end{aligned}$$

Now, suppose $m(i) = 1$. Let (x, y) be any point such that $x_{i-1} < x < x_i$ and $y = f(x)$. Further, let $F(x) - F(x_{i-1}) = r$. Observe that $G(y) - G(y_{\sigma(i)-1}) = r$. Then,

$$\phi(x, y) = (F(x), G(y)) = (F(x_{i-1}) + r, G(y_{\sigma(i)-1}) + r).$$

It now follows that the mass which C deposits in the square $[F(x_{i-1}), F(x_i)] \times [G(y_{\sigma(i)-1}), G(y_{\sigma(i)})]$ is spread uniformly along the diagonal from $(F(x_{i-1}), G(y_{\sigma(i)-1})$ to $(F(x_i), G(y_{\sigma(i)}))$. If $m(i) = -1$, the proof is similar that the mass which C deposits in the square $[F(x_{i-1}), F(x_i)] \times [G(y_{\sigma(i)-1}), G(y_{\sigma(i)})]$ is spread uniformly along the diagonal from $(F(x_{i-1}), G(y_{\sigma(i)}))$ to $(F(x_i), G(y_{\sigma(i)-1}))$. This proves that C is the shuffle of Min generated by $(n, \{F(x_i)\}, \{G(y_i)\}, \sigma, m)$.

Remark: The proofs of Theorem 2.1 and 2.2 given here are modifications of the proof of Fréchet's probabilistic interpretation of M given by E.F. Wolff in [10].

Theorems 2.1 and 2.2 yield at once the following corollaries:

Corollary 2.1. The random vector (X, Y) has a shuffle of Min as its connecting copula if and only if each of X and Y is almost surely a strongly piecewise monotone function of the other.

Corollary 2.2. The random vector (X, Y) has a straight shuffle of Min as its connecting copula if and only if each of X and Y is almost surely a strongly piecewise nonincreasing function of the other.

Corollary 2.3. The random vector (X, Y) has a flipped shuffle of Min as its connecting copula if and only if each of X and Y is almost surely a strongly piecewise non increasing function of the other.

Remark. If $Y = f \circ X$ almost surely then (because G is continuous) X lies with probability zero in any interval on which f is constant and there are at most countably many such intervals. Thus, if one is willing to relax the condition that the domain of f is $\overline{\mathbb{R}}$, the function f which relates X and Y in each of the preceding theorems and corollaries may be chosen to be invertible.

3. Dense families in COP.

G. Kimeldorf and A.R. Sampson [4] prove essentially that the copula for independence, π, can be approximated by certain shuffles of Min. The next theorem shows that any copula can be so approximated.

Theorem 3.1. Straight (flipped) shuffles of Min are dense in COP endowed with the sup norm.

Proof: Let C be an arbitrary copula and let $\epsilon > 0$ be given. Since the proofs for straight and flipped shuffles of Min are so similar, we give only the first by constructing C^*, a straight shuffle of Min, such $\|C - C^*\| < \epsilon$ where $\|\cdot\|$ is the sup norm. Using (1.3), choose a positive integer K such that for any copula C',

$$|C'(x_1, y_1) - C'(x_0, y_0)| < \epsilon/2 \quad \text{whenever} \tag{3.1}$$

$$|x_1 - x_0| < 1/K \quad \text{and} \quad |y_1 - y_0| < 1/K.$$

Next subdivide I^2 into K vertical columns and K horizontal rows as follows:

$$V_i = [(i-1)/K, i/K] \times I \quad \text{and} \quad H_j = I \times [(j-1)/K, j/K]$$

for $i, j = 1, 2, \ldots, K$. Set $S_{ij} = V_i \cap H_j$ and let $m_{ij} = \mu_C(S_{ij})$. Since μ_C is a dsm,

$$\mu_C(V_i) = m_{i1} + m_{i2} + \ldots + m_{iK} = 1/K$$

and

$$\mu_C(H_j) = m_{1j} + m_{2j} + \ldots + m_{Kj} = 1/K.$$

Subdivide each V_i into K vertical subcolumns, labeled from left to right $V_{i1}, V_{i2}, \ldots, V_{iK}$, so that the width of V_{ik} is m_{ik}. Similarly subdivide each H_j into K horizontal subrows, labeled from bottom to top $H_{j1}, H_{j2}, \ldots, H_{jK}$, so that the height of H_{jk} is m_{kj}. Then $V_{ij}\ H_{ji}$ is a square with sides of length m_{ij} located in S_{ij}. For each $i, j = 1, 2, \ldots, K$, spread a mass of size m_{ij} uniformly along the diagonal of $V_{ij}\ H_{ji}$ which has positive slope. This is clearly a mass distribution for a straight shuffle of Min which we denote by C^*. Since $\mu_{C^*}(S_{ij}) = m_{ij} = \mu_C(S_{ij})$ we have $C(i/K, j/K) = C^*(i/K, j/K)$ for $i, j = 0, 1, \ldots, K$. Finally, let $(x, y) \in I^2$. There is some $i, j = 0, 1, \ldots,\ K$ such that

$|x - i/K| < 1/K$ and $|y - j/K| < 1/K$. Thus, by (3.1) we have

$$\begin{aligned} |C(x,y) - C^*(x,y)| &\le |C(x,y) - C(i/k, j/K)| + |C(i/K, j/K) - C^*(i/K, j/K)| \\ &\quad + |C^*(i/K, j/K) - C^*(x,y)| \\ &< \frac{\epsilon}{2} + 0 + \frac{\epsilon}{2} = \epsilon, \end{aligned}$$

and the proof is complete.

4. Approximation.

The usual understanding of two random variables being stochastically independent is that knowledge of one variable's value gives no helpful information regarding the value of the other. On the other hand, if one random variable is a function of the other, then knowing the value of one variable gives complete information regarding the value of the other. Nevertheless we have the following theorem which is interesting even in the case of stochastic dependence.

Theorem 4.1. Given any $\epsilon > 0$ and any pair of continuously distributed random variables X, Y, there exist X^* and Y^* and an invertible piecewise increasing function f such that $Y^* = f \circ X^*$, $F = F^*$, $G = G^*$, and $\|H - H^*\| < \epsilon$.

Proof: Let C be the connecting copula for (X, Y). Let C' be the straight shuffle of Min which uniformly approximates C within ϵ, i.e.,

$$|C(s,t) - C'(s,t)| < \epsilon \quad \text{whenever } s, t \text{ are in } I. \tag{4.1}$$

The function H' defined on $\mathbb{R}^2$ via

$$H'(x,y) = C'(F(x), G(y)) \tag{4.2}$$

is a two-dimensional distribution function, see [6]. Let X' and Y' be the orthogonal projections of $\mathbb{R}^2$ onto the $x-$ and $y-$axes, respectively where $\mathbb{R}^2$ is endowed with the Lebesgue-Stieltjes measure induced by H'. Clearly H' is the distribution function of (X', Y'), $F' = F$

and $G' = G$. Moreover, by Corollary 2.2 and the Remark following Corollary 2.3, there is an invertible piecewise increasing function f such that $Y' = f \circ X'$ almost surely. It follows from (4.1) and (4.2) that for any x, y in $\mathbf{R}$,

$$|H(x,y) - H'(x,y)| = |C(F(x), G(y)) - C'(F(x), G(y))| < \epsilon.$$

Finally let $X^* = X'$ and $Y^* = f \circ X^*$ to complete the proof.

As this manuscript was being prepared for publication, we learned via a private communication from R.A. Vitale that he also discovered the preceding result form a different perspective. The reader is encouraged to look at Vitale's paper [9]; his formulation of the result will be especially useful for simulation.

In 1959, A. Rényi [5] proposed the following set of axioms for a measure of dependence $R(X,Y)$ for pairs of random variables (X,Y) which were not restricted to being continuously distributed:

(A) $R(X,Y)$ is defined for any X and Y.

(B) $R(X,Y) = R(Y,X)$.

(C) $0 \leq R(X,Y) \leq 1$.

(D) $R(X,Y) = 0$ if and only if X and Y are independent.

(E) $R(X,Y) = 1$ if either $X = f(Y)$ or $Y = g(X)$ for some Borel- measurable functions f and g.

(F) If f and g are Borel-measurable one-one mappings of the real line into itself, then $R(f(X), g(Y)) = R(X,Y)$.

(G) If the joint distribution of X and Y is bivariate normal, with correlation coefficient r, then $R(X,Y) = |r|$.

In 1981 Schweizer and Wolff [7] used copulas to define several natural nonparametric measures of dependence of pairs of continuously distributed random variables and then showed that these measures satisfy (A)-(D) and (E')- (H') where

(E') $R(X, Y) = 1$ if and only if each of X and Y is almost surely a strictly monotone function of the other.

(F') If f and g are strictly monotone almost surely on Range X and Range Y, respectively, then $R(f(X), g(Y)) = R(X, Y)$.

(G') If the joint distribution of X and Y is bivariate normal, with correlation coefficient r, then $R(X, Y)$ is a strictly increasing functions ϕ of $|r|$.

(H') If (X, Y) and (X_n, Y_n), $n = 1, 2, \ldots$, are pairs of random variables with joint distribution functions H and H_n, respectively, and if the sequence $\{H_n\}$ converges weakly to H, then $\lim_{n\to\infty} R(X_n, Y_n) = R(X, Y)$.

In light of Theorem 4.1 and in the presence of (A) and (D), one must choose between axioms (E) and (H'); they cannot both be true. To see this, let X and Y be independent normally distributed random variables. We may by Theorem 4.1 construct a sequence (X_n, Y_n) such that $Y_n = f_n(X_n)$ and $\{H_n\}$ converges weakly to H. If the measure of dependence satisfies both (D) and (E), then $R(X, Y) = 0$ while, for each n, $R(X_n, Y_n) = 1$. This contradicts (H').

If a measure of dependence satisfies (A), (D) and (E'), then at most one of the axioms (F) and (H') can be true. To see this, let X, Y, X_n, Y_n and f_n be as in the preceding paragraph. A close look at the proof of Theorem 2.1 shows that f_n can be chosen to have domain $\mathbf{R}$ and still be a one-one function in this case because $F_n^{-1}(\{s\}) \times G_n^{-1}(\{t\})$ is a singleton for each s, t in I. By (E'), for each n, we have $R(X_n, X_n) = 1$. If (F) is true, then $R(X_n, Y_n) = R(X_n, f_n(X_n)) = R(X_n, X_n) = 1$. But since $R(X, Y) = 0$, this again contradicts (H').

ACKNOWLEDGEMENTS

The authors wish to thank B. Schweizer for helpful suggestions which improved this paper. Thanks also to R.B. Nelsen for bringing some related references to our attention.

References.

[1] Brown, J.R., (1966), Approximation Theorems for Markov operators. *Pacific J. Math.* 16, 13-23.

[2] Fréchet, M., (1951), Sur les tableaux de correlation dont les marges sont données. *Ann. Univ. Lyon Sect. A.* 14, 53-77.

[3] Kamiński, A., Mikusiński, P., Sherwood, H. and Taylor, M.D., Doubly stochastic measures, topology and latticework hairpins. (To appear in *J. Math. Anal. Appl.*)

[4] Kimeldorf, G. and Sampson, A.R., (1978), Monotone dependence. *Ann. of Stat.* Vol. 6, N. 4, 895-903.

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

[6] Schweizer, B. and Sklar, A., (1983), *Probabilistic Metric Spaces*, New York: North-Holland.

[7] Schweizer, B. and Wolff, E.F., (1981), On nonparametric measures of depedence for random variables. *Ann. of Stat.* Vol. 9, N. 4, 879-885.

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

[9] Vitale, R.A., Stochastic dependence on an interesting class of degenerate distributions. (to appear)

[10] Wolff, E.F., (1977), Measures of dependence derived from copulas, Ph.D. Thesis, Univ. Massachusetts, Amherst.

Dept. of Mathematics
University of Central Florida
Orlando, Florida 32816. US