Internat. J. Math. & Math. Sci.
VOL. 16 NO. 1 (1993) 155-164

THE FRÉCHET TRANSFORM

PIOTR MIKUSIŃSKI, MORGAN PHILLIPS
and
HOWARD SHERWOOD, MICHAEL D. TAYLOR

Department of Mathematics
University of Central Florida
Orlando, FL 32816

(Received April 19, 1991 and in revised form July 7, 1992)

ABSTRACT. Let $F_1, \cdots, F_N$ be 1-dimensional probability distribution functions and C be an N-copula. Define an N-dimensional probability distribution function G by $G(x_1, \cdots, x_N) = C(F_1(x_1), \cdots, F_N(x_N))$. Let ν be the probability measure induced on $\mathbb{R}^N$ by G and μ be the probability measure induced on $[0,1]^N$ by C. We construct a certain transformation Φ of subsets of $\mathbb{R}^N$ to subsets of $[0,1]^N$ which we call the Fréchet transform and prove that it is measure-preserving. It is intended that this transform be used as a tool to study the types of dependence which can exist between pairs or N-tuples of random variables, but no applications are presented in this paper.

KEY WORDS AND PHRASES. Fréchet transform, probability distribution function, doubly stochastic measure, copula, distributions with fixed marginals.
1991 Mathematics Subject Classification: 28A35, 60A10

1. INTRODUCTION.

The genesis of this work lies in attempts to analyze the types of dependence which can exist between pairs of random variables. To explain this statement we must first introduce a number of terms.

By a <u>2-copula</u> we mean a function $C:[0,1]^2 \to [0,1]$ satisfying

$$C(a,0) = C(0,a) = 0,$$

$$C(a,1) = C(1,a) = a,$$

and $\quad C(a,b) - C(c,b) - C(a,d) + C(c,d) \geq 0$ whenever $a \geq c$ and $b \geq d$.

By a <u>doubly stochastic measure</u> on $[0,1]^2$ we mean a probability measure μ defined on (at least) the Borel sets of $[0,1]^2$ and satisfying

$$\mu(A \times [0,1]) = \mu([0,1] \times A) = \lambda(A)$$

for all Lebesgue measurable subsets A of $[0,1]$, λ being 1-dimensional Lebesgue measure.

It is well-known that there is a one-to-one correspondence between 2-copulas and doubly stochastic measures, the correspondence being defined by the equation

$$C(a,b) = \mu([0,a) \times [0,b)).$$

Let X and Y be random variables with distribution functions F and G respectively and joint distribution function H. It is known (see [1] or [2]) that a copula C can always be found which satisfies

$$H(x,y) = C(F(x), G(y))$$

for all real numbers x and y. If F and G are continuous, then C is uniquely determined, otherwise it is not. Because C connects the joint distribution function to its marginals, one may identify C (or its corresponding doubly stochastic measure) with the type of dependence which exists between the random variables X and Y. For example (see [3]) X and Y are independent if and only if C can be chosen to be $C(a,b) = a \cdot b$, and Y is a nondecreasing function of X if and only if C can be chosen to be $C(a,b) = \min(a,b)$.

In [4] and [5], three of the authors of this paper have investigated the question of characterizing copulas that correspond to types of dependence between pairs of random variables different from those described above. A useful tool in those investigations was the map $\phi: \mathbb{R}^2 \rightarrow [0,1]^2$ defined by $\phi(x,y) = (F(x), G(y))$. Loosely speaking this map takes horizontal lines to horizontal lines, vertical lines to vertical lines, and preserves the relations of above/below and right/left. Also, if ν is the probability measure induced on $\mathbb{R}^2$ by the joint distribution function H, then ϕ induces a probability measure μ on $[0,1]^2$ via the equation $\mu(A) = \nu(\phi^{-1}(A))$ for all A such that $\phi^{-1}(A)$ is ν-measurable. Provided F and G are both continuous, μ is a doubly stochastic measure and defines the copula C for X and Y and hence the type of dependence for this pair of random variables. It is both straightforward and illuminating to see how this idea can be used to show that Y is a nondecreasing function of X if and only if $C = \min$.

However if either F or G has discontinuities, then μ will have positive mass clumped at points or along horizontal or vertical line segments and will fail to be doubly stochastic. To see what is meant by "clumping", consider a simple, 1-dimensional example. Let ν be the probability measure on $\mathbb{R}$ with distribution function

$$F(x) = \begin{cases} \frac{1}{2}\left(\frac{1}{1+e^{-x}}\right) & \text{for } x \leq 0 \\ \frac{1}{2}\left(1 + \frac{1}{1+e^{-x}}\right) & \text{for } x > 0. \end{cases}$$

We would like to think of F as inducing a probability measure μ on [0,1] via the relation $\mu(A) = \nu(F^{-1}(A))$ whenever $F^{-1}(A)$ is ν-measurable. It is easily seen that $\mu([0,\frac{3}{4}]) = \frac{1}{4}$, however $\mu([0,\frac{3}{4}+\epsilon]) = \frac{3}{4}$ for arbitrarily small positive ϵ. Thus, in some sense, the discontinuity of F at 0 has caused a mass of 1/2 to clump at the point 3/4 in [0,1].

This clumping in turn makes it difficult and complicated to show that a certain copula corresponds to a certain type of dependence between a pair of random variables. Fréchet in [3] gave the original proof that the copula min corresponds to having Y almost surely a nondecreasing function of X and the copula $W(x,y) = \max(x+y-1,0)$ corresponds to having Y almost surely a nonincreasing function of X. He contented himself with giving the proof in the case when the distribution functions of X and Y were continuous and claiming the general case clearly followed, a very believable claim. One of us was able to construct a rigorous and general proof Fréchet's theorem using ϕ and found that the proof took eighty pages. (A complete proof along totally different lines is given in [6].)

The purpose of this paper is to replace ϕ by a map which does essentially the same things but does not cause mass to "clump up" in $[0,1]^2$ and destroy the doubly stochastic character of probability measures which one seeks to induce in the unit square. In recognition of the key role played by Fréchet's work in stimulating both this and earlier work on our part, we call the new map the Fréchet transform. It is hoped that this new map can ultimately be used to simplify the analysis of types of dependence between random variables. The Fréchet transform Φ, which we shall describe presently, is a set-transformation rather than a point-transformation, and we describe it in a setting which is appropriate for considering N random variables, not just two. In this setting, if $F_1, \cdots, F_N$ are the distribution functions of the random variables and H is the joint distribution function, then by Sklar's theorem (see [1] or [2]) one can always find an N-copula (to be defined later) C which satisfies

$$H(x_1, \cdots, x_N) = C(F_1(x_1), \cdots, F_N(x_N)).$$

If ν is the probability measure induced by H and μ is the probability measure induced on the unit N-cube by C, then our main result is that $\mu(\Phi(A)) = \nu(A)$ whenever A is ν-measurable.

Some of this work is drawn from [7].

We are much indebted to the referee for his insightful and helpful comments.

2. THE FRÉCHET TRANSFORM FOR 1-DIMENSIONAL DISTRIBUTION FUNCTIONS.

By I we mean the unit interval $[0,1]$ and by $\overline{\mathbb{R}}$ we mean the closed reals, $[-\infty, +\infty]$. Let $F: \overline{\mathbb{R}} \to I$ be a nondecreasing function satisfying $F(-\infty) = 0$ and $F(\infty) = 1$. Recall that 2^A stands for the collection of subsets of A.

DEFINITION 1. We define $\Phi: 2^{\overline{\mathbb{R}}} \to 2^I$, the Fréchet transform determined by F, by

$$\Phi(A) = \{y \in [0,1] : \text{there is an } x \in A \text{ such that } F(x^-) \le y \le F(x^+)\}.$$

We take $F(-\infty^-)$ to be $F(-\infty)$ and $F(\infty^+)$ to be $F(\infty)$.

To see what the Fréchet transform does, picture the graph of F lying in the xy-plane in the infinite strip $-\infty \le x \le \infty$, $0 \le y \le 1$. Wherever the graph of F is broken by discontinuities, we fill in the breaks with vertical line segments which stretch from $F(x^-)$ to $F(x^+)$. Call the graph of F with these line segments added F^*. F^* is a continuous curve running from $x = -\infty$ to $x = \infty$ and rising from $y = 0$ to $y = 1$. Now think of A as a set lying in the x-axis below F^*. Project A upward onto F^*. Whenever a point x of A corresponds to a discontinuity of F, the projection of x upwards is thought of as being "smeared" over the whole vertical line segment in F^* which lies above x. Denote the resulting set, this "projection" of A, by the symbol P. Now let P be projected horizontally onto the y-axis. This projection, a set lying in the unit interval, is what we mean by $\Phi(A)$. See Figures 1, 2, and 3 for an illustration of this situation.

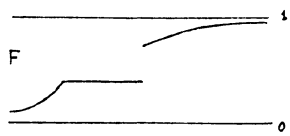

Figure 1

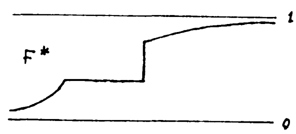

Figure 2

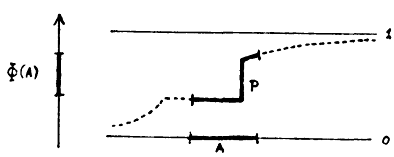

Figure 3

NOTE: The application of Φ to a singleton set plays such an important role here that we shall feel free to write $\Phi(p)$ or $\Phi(x)$ when we really mean $\Phi(\{p\})$ or $\Phi(\{x\})$.

The proofs of the next three results are straightforward and are omitted.

PROPOSITION 1. For every $p \in \overline{\mathbb{R}}$ we have $\Phi(p) = [F(p^-), F(p^+)]$.

See Figure 4 for an illustration.

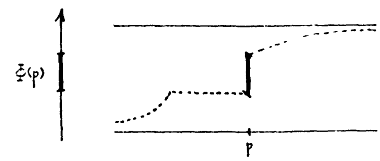

Figure 4

PROPOSITION 2. Φ distributes over unions.

Notice this implies the useful facts that $\Phi(A) \subseteq \Phi(B)$ whenever $A \subseteq B$ and that $\Phi(A) = \bigcup\{\Phi(x): x \in A\}$.

DEFINITION 2. If A is a subset of $\overline{\mathbb{R}}$, then

$$A^* = \cup\{S: \Phi(S) = \Phi(A)\}.$$

This means $\Phi(A^*) = \Phi(A)$ and A^* is the maximal set having this property. Geometrically the process of forming A^* from A amounts to the following: Whenever A contains a point x such that F is constant over an interval J containing x, then we "add" J to A. A^* amounts to the union of A and all such J. See Figures 5 and 6. Note also that if $A \subseteq B$, then $A^* \subseteq B^*$.

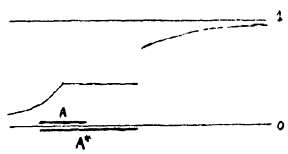

Figure 5

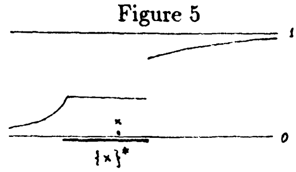

Figure 6

PROPOSITION 3. Let $\mathcal{D} = \{A^*: A \subseteq \overline{\mathbb{R}}\}$ and $\mathcal{R} = \{\Phi(A): A \subseteq \overline{\mathbb{R}}\}$. Then Φ acts as a one-to-one map of $\mathcal{D}$ onto $\mathcal{R}$. We will designate the inverse of this map from $\mathcal{D}$ to $\mathcal{R}$ as Ψ.

PROPOSITION 4. For every $A \subseteq \overline{\mathbb{R}}$ and every $p \in \overline{\mathbb{R}}$, we have $p \in A^*$ if and only if there is an $x \in A$ such that $\Phi(p) \subseteq \Phi(x)$.

PROOF. Suppose $p \in A^*$. We must have $\Phi(p) \subseteq \Phi(A^*) = \Phi(A)$. We know that $\Phi(A) = \cup\{\Phi(x): x \in A\}$. If $\Phi(p)$ is a singleton, then there must be some $x \in A$ such that $\Phi(p) \subseteq \Phi(x)$ and we are done. Suppose $\Phi(p)$ is not a singleton. Recall that $\Phi(p) = [F(p^-), F(p^+)]$. It must be possible to find y such that $F(p^-) < y < F(p^+)$. Since $y \in \Phi(p) \subseteq \Phi(A)$, there must be some $x \in A$ such that $y \in \Phi(x)$, i.e., such that $F(x^-) \le y \le F(x^+)$. If $x < p$, then we must have $F(x^+) \le F(p^-) < y$ which is impossible. If $p < x$, then we must have $y < F(p^+) \le F(x^-)$ which is equally impossible. It follows that $p = x \in A$ so that $\Phi(p) \subseteq \Phi(x)$ trivially.

Now suppose there is an $x \in A$ such that $\Phi(p) \subseteq \Phi(x)$. We must have $\Phi(A^*) \subseteq \Phi(\{p\} \cup A^*) = \Phi(p) \cup \Phi(A^*) \subseteq \Phi(x) \cup \Phi(A^*) = \Phi(\{x\} \cup A^*) = \Phi(A^*)$. So $\Phi(\{p\} \cup A^*) = \Phi(A^*)$, and by the maximal nature of A^* it follows that $p \in A^*$. □

PROPOSITION 5. Let $\{A_\alpha\}$ be an indexed family of subsets of $\bar{\mathbf{R}}$. Then $\cup_\alpha A_\alpha{}^* = (\cup_\alpha A_\alpha)^*$. (This is equivalent to saying $\mathfrak{D}$ is closed under arbitrary unions.)

PROOF. Let $p \in (\cup_\alpha A_\alpha)^*$. There must exist $x \in \cup_\alpha A_\alpha$ such that $\Phi(p) \subseteq \Phi(x)$. Then there must be β such that $x \in A_\beta$, and hence $p \in A_\beta{}^*$. Thus $(\cup_\alpha A_\alpha)^* \subseteq \cup_\alpha A_\alpha{}^*$. Containment in the other direction follows trivially. □

NOTE: The referee has drawn our attention to the fact that $*$ is a topological closure operation. The observation is intriguing, but its significance is not yet clear.

PROPOSITION 6. Ψ distributes over unions.

PROOF. Let $\{B_\alpha\}$ be a family of sets such that each B_α is in the domain of Ψ. For each α we can find a set A_α such that $\Phi(A_\alpha{}^*) = B_\alpha$. Then $\cup_\alpha B_\alpha = \cup_\alpha \Phi(A_\alpha{}^*) = \Phi(\cup_\alpha A_\alpha{}^*) = \Phi((\cup_\alpha A_\alpha)^*)$ which is in the domain of Ψ. Therefore $\Psi(\cup_\alpha B_\alpha) = \Psi(\cup_\alpha \Phi\Psi(B_\alpha)) = \Psi\Phi(\cup_\alpha \Psi(B_\alpha)) = \cup_\alpha \Psi(B_\alpha)$. □

We now show that Ψ admits a very nice characterization. The key concept arises from the following considerations:

Suppose $p < q$. We must have $F(p^+) \leq F(q^-)$. From this it follows that $\Phi(p) = [F(p^-), F(p^+)]$ and $\Phi(q) = [F(q^-), F(q^+)]$ can have in common at most an endpoint.

Recall how, when the definition of Fréchet transform was first introduced, we considered F^*, a curve obtained by adding to the graph of F the vertical line segments corresponding to the discontinuities of F. We may think of $\Phi(p)$ and $\Phi(q)$ as corresponding to $\{p\} \times \Phi(p)$ and $\{q\} \times \Phi(q)$. These are vertical line segments in F^*, possibly degenerate ones.

To say that

$$F(p^-) \leq F(q^-) \leq F(q^+) \leq F(p^+)$$

is to say that the line segment $\{q\} \times \Phi(q)$ is a subset of the line segment $\{p\} \times \Phi(p)$. This can happen if and only if

$$F(p^-) \leq \frac{F(q^-)+F(q^+)}{2} \leq F(p^+).$$

PROPOSITION 7. Define

$$\widetilde{F}(x) = \frac{F(x^-) + F(x^+)}{2}.$$

Then for all $B \in \mathfrak{D}$ we have $\widetilde{F}^{-1}(B) = \Psi(B)$.

PROOF. Since B must have the form $\Phi(A^*)$ and $\Phi, \Psi, \widetilde{F}^{-1}$, and $*$ all distribute over unions, it is sufficient to prove the result for sets B of the form $\Phi(\{p\}^*)$. Further, since $\Psi(\Phi(\{p\}^*)) = \{p\}^*$, we need only show $\widetilde{F}^{-1}(\Phi(\{p\}^*)) = \{p\}^*$. The following statements are equivalent.

$q \in \widetilde{F}^{-1}(\Phi(\{p\}^*))$.

$\widetilde{F}(q) \in \Phi(\{p\}^*)$.

$\widetilde{F}(q) \in \Phi(p)$.

$F(p^-) \leq \widetilde{F}(q) \leq F(p^+)$.

$F(p^-) \leq \dfrac{F(q^-) + F(q^+)}{2} \leq F(p^+)$.

$F(p^-) \leq F(q^-) \leq F(q^+) \leq F(p^+)$.

$\Phi(q) \subseteq \Phi(p)$.

$q \in \{p\}^*$. □

3. QUASI-INVERSES AND PROBABILITY MEASURES; THE 1-DIMENSIONAL CASE.

It is the object of this section to prove that the Fréchet transform Φ, and hence its inverse Ψ, are measure preserving. Throughout this section we shall assume all the cumulative probability distribution functions with which we deal are left-continuous.

By λ we mean 1-dimensional Lebesgue measure.

If H is a 1-dimensional distribution function, by the quasi-inverse of H we mean the function $\widehat{H}: I \rightarrow \overline{\mathbf{R}}$ defined by $\widehat{H}(x) = sup\{t: H(t) < x\}$. $\widehat{H}$ is left-continuous. In the discussions below it is helpful to keep the following in mind:

If $t < \widehat{H}(x)$, then $H(t) < x$.

If $t > \widehat{H}(x)$, then $H(t) \geq x$.

Indeed these properties characterize the quasi-inverse. The concept of a quasi-inverse generalizes the idea of an inverse of a function, at least in the case of continuous functions. An extensive discussion of quasi-inverses is given in [1].

Let F be a fixed, 1-dimensional distribution function and let

ν = the probability measure induced on $\overline{\mathbf{R}}$ by the distribution function F,

$D = \{y \in I: F$ has the value y over some nondegenerate interval $\}$.

The set D is at most countable, so $\lambda(D) = 0$. We shall use this fact shortly. Notice also that $\widehat{F}^{-1}$ maps subsets of $\overline{\mathbf{R}}$ to subsets of I the same as Φ does. These two maps are almost the same in a sense which we now make precise.

PROPOSITION 8. For every $A \subseteq \overline{\mathbf{R}}$ we have $\widehat{F}^{-1}(A) \subseteq \Phi(A) \subseteq \widehat{F}^{-1}(A) \cup D$.

PROOF. Let $y \in \widehat{F}^{-1}(A)$ and set $x = \widehat{F}(y)$. Note that $x \in A$. Notice that if $t < x = \widehat{F}(y)$, then $F(t) < y$. From this we deduce $F(x) \leq y$. If, on the other hand, $t > x = \widehat{F}(y)$, then $F(t) \geq y$. From this we deduce $F(x^+) \geq y$. Thus there is an $x \in A$ with the property that $F(x) \leq y \leq F(x^+)$. Thus $y \in \Phi(A)$.

Now suppose $y \in \Phi(A)$ and $y \notin \widehat{F}^{-1}(A)$. We need only show $y \in D$ and we are done. There must be an $x \in A$ such that $F(x) \leq y \leq F(x^+)$. Notice that $\widehat{F}(y) \notin A$ so that $x \neq \widehat{F}(y)$. Suppose $x < \widehat{F}(y)$. We must have $F(x) < y \leq F(x^+)$ by the definition of quasi-inverse. But consider u lying in the interval $x < u < \widehat{F}(y)$. By the definition of quasi-inverse we have $F(u) < y$, but since F is nondecreasing we must also have $F(x^+) \leq F(u) < y$ which is a contradiction. Therefore we must have $x > \widehat{F}(y)$. Now consider u lying in the interval $\widehat{F}(y) < u < x$. We must have $F(x) \geq F(u) \geq y \geq F(x)$ which means F takes on the constant value y on the interval $\widehat{F}(y) < u < x$. Thus $y \in D$. □

THEOREM; 1-DIMENSIONAL VERSION. If A is ν-measurable, then $\Phi(A)$ is Lebesgue measurable and $\mu[\Phi(A)] = \nu(A)$. If $B \in \mathfrak{R}$ and B is Lebesgue measurable, then $\Psi(B)$ is ν-measurable and $\nu[\Psi(B)] = \mu(B)$.

PROOF. We first consider the special case of a half-open interval $[a,b)$ in $\overline{\mathbf{R}}$ and show that its measure is preserved under the transformation Φ.

We know that $\widehat{F}^{-1}([a,b)) \subseteq \Phi([a,b)) \subseteq \widehat{F}^{-1}([a,b)) \cup D$. Since $\lambda(D) = 0$, it follows that $\lambda[\Phi([a,b))] = \lambda[\widehat{F}^{-1}([a,b))]$. Let $B = \widehat{F}^{-1}([a,b))$. We will show B must be an interval. Choose u and v from B such that $u < v$. We see that $a \leq \widehat{F}(u) \leq \widehat{F}(v) < b$. Suppose $u < w < v$. We must have $a \leq \widehat{F}(u) \leq \widehat{F}(w) \leq \widehat{F}(v) < b$. Thus $\widehat{F}(w) \in [a,b)$, which means $w \in B$. Therefore B is an interval.

We need to show $\lambda(B) = F(b) - F(a)$ in order to show Φ preserves measure when applied to the half-open interval $[a,b)$. It might at this point be helpful to recall and somewhat expand upon the important properties of the quasi-inverse.

$u < \widehat{F}(v)$ implies $F(u) < v$.
$u > \widehat{F}(v)$ implies $F(u) \geq v$.
$u = \widehat{F}(v)$ implies $F(u) \leq v$.

Also

$F(u) < v$ implies $u \leq \widehat{F}(v)$.
$F(u) > v$ implies $u > \widehat{F}(v)$.
$F(u) = v$ implies $u \geq \widehat{F}(v)$.

The second set of properties follows readily from the first set. Now let c and d be the left-hand and right-hand endpoints respectively of B. If $c < t <$ b, then we successively deduce the following:

$\widehat{F}(t) \in [a,b)$.
$a \leq \widehat{F}(t) < b$.
$F(a) \leq t \leq F(b)$.

Hence $(c,d) \subseteq [F(a), F(b)]$. Now let us suppose we have t such that $F(a) < t < F(b)$. We deduce the following:

$a \leq \widehat{F}(t) < b$.
$\widehat{F}(t) \in [a,b)$.
$t \in B$.

Hence $(F(a), F(b)) \subseteq B \subseteq [c,d]$. It follows from these containments that $\lambda(B) = d - c = F(b) - F(a) = \nu([a,b))$. Thus Φ is measure preserving when applied to $[a,b)$.

Now let A be a ν-measurable subset of $\overline{\mathbf{R}}$. Since $\widehat{F}$ is a measurable function and D has Lebesgue measure zero, it follows from Proposition 8 that $\Phi(A)$ is Lebesgue measurable. Choose $\epsilon > 0$. There must exist a sequence of intervals $\{I_n\}$, each I_n of the form $[a_n, b_n)$, such that $A \subseteq \cup_n I_n$ and

$$\sum_n \nu(I_n) \leq \nu(A) + \epsilon.$$

We see from this that

$$\begin{aligned}\lambda[\Phi(A)] &\leq \lambda[\Phi(\cup_n I_n)]\\ &= \lambda[\cup_n \Phi(I_n)]\\ &\leq \sum_n \lambda[\Phi(I_n)]\\ &= \sum_n \nu(I_n)\\ &\leq \nu(A) + \epsilon.\end{aligned}$$

Since ϵ is arbitrary, we must have $\lambda[\Phi(A)] \leq \nu(A)$. Similarly $\lambda[\Phi(A^c)] \leq \nu(A^c)$ where A^c is the complement of A. We then see that

$$1 = \lambda[\Phi(\overline{\mathbf{R}})] \leq \lambda[\Phi(A)] + \lambda[\Phi(A^c)] \leq \nu(A) + \nu(A^c) = 1$$

which can only be true if $\lambda[\Phi(A)] = \nu(A)$. Thus Φ is measure preserving.

Now suppose $B \in \mathfrak{R}$ and B is Lebesgue measurable. The ν-measurability of $\Psi(B)$ follows from Proposition 7. Finally $\lambda(B) = \lambda[\Phi\Psi(B)] = \nu[\Psi(B)]$. □

4. THE FRÉCHET TRANSFORM FOR DIMENSION N.

We now extend our results to higher dimensions. It is straightforward to see that almost all concepts and operations can be defined or applied in a componentwise fashion. The proofs for the 1-dimensional case carry over to the N-dimensional case for the most part with little or no change. Because of these considerations, we shall not bother to prove most of the propositons in this section.

For $i = 1, 2, \cdots, N$, we take $F_i: \overline{\mathbf{R}} \to [0,1]$ be a nondecreasing function satisfying $F_i(-\infty) = 0$ and $F_i(\infty) = 1$. We take $F_i(-\infty^-)$ to be $F_i(-\infty)$ and $F_i(\infty^+)$ to be $F_i(\infty)$. Let $\Phi_i: 2^{\overline{\mathbf{R}}} \to 2^{[0,1]}$ be the Fréchet transform determined by F_i for $i =$ 1 to N.

DEFINITION 3. We define $\Phi: 2^{\overline{\mathbb{R}}^N} \to 2^{I^N}$, the Fréchet transform determined by $(F_1, F_2, \cdots, F_N)$, by

$\Phi(A) = \{(y_1, \cdots, y_N) \in I^N$: there exists $(x_1, \cdots, x_N) \in A$ such that $y_i \in \Phi_i(x_i)$ for $i = 1$ to $N\}$ where A is an arbitrary subset of $\mathbb{R}^N$.

DEFINITION 4. If A is a subset of $\overline{\mathbb{R}}^N$, then

$$A^* = \cup \{S: \Phi(S) = \Phi(A)\}.$$

(This means $\Phi(A^*) = \Phi(A)$ and A^* is the maximal set having this property.)

PROPOSITION 9. For every $p = (p_1, \cdots, p_N) \in \overline{\mathbb{R}}$ we have

$$\Phi(p) = [F_1(p_1^-), F_1(p_1^+)] \times \cdots \times [F_N(p_N^-), F_N(p_N^+)] = \Phi_1(p_1) \times \cdots \times \Phi_N(p_N).$$

PROPOSITION 10. Φ distributes over unions.

PROPOSITION 11. Let $\mathfrak{D}_N = \{A^*: A \subseteq \overline{\mathbb{R}}^N\}$ and $\mathfrak{R}_N = \{\Phi(A): A \subseteq \overline{\mathbb{R}}^N\}$. Then Φ acts as a one-to-one map of $\mathfrak{D}_N$ onto $\mathfrak{R}_N$. We will designate the inverse of this map from $\mathfrak{D}_N$ to $\mathfrak{R}_N$ as Ψ.

PROPOSITION 12. For every $A \subseteq \overline{\mathbb{R}}$ and every $p \in \overline{\mathbb{R}}$, we have $p \in A^*$ if and only if there is an $x \in A$ such that $\Phi(p) \subseteq \Phi(x)$.

PROPOSITION 13. Let $\{A_\alpha\}$ be an indexed family of subsets of $\overline{\mathbb{R}}^N$. Then $\cup_\alpha A_\alpha^* = (\cup_\alpha A_\alpha)^*$. (This is equivalent to saying $\mathfrak{D}_N$ is closed under arbitrary unions.)

PROPOSITION 14. Ψ distributes over unions.

We also give a couple of results peculiar to the N-dimensional case. The proofs are straightforward.

PROPOSITION 15. $\Phi(A_1 \times \cdots \times A_N) = \Phi_1(A_1) \times \cdots \times \Phi_N(A_N)$.

PROPOSITION 16. $(A_1 \times \cdots \times A_N)^* = A_1^* \times \cdots \times A_N^*$.

We note that the characterization we gave of Ψ can be extended to the N-dimensional case. For $i = 1$ to N let

$$\widetilde{F}_i(x) = \frac{F_i(x^-) + F_i(x^+)}{2}$$

and set

$$\widetilde{F}(x_1, \cdots, x_N) = (\widetilde{F}_1(x_1), \cdots, \widetilde{F}_N(x_N)).$$

PROPOSITION 17. For every $B \in \mathfrak{D}_N$ we have $\widetilde{F}^{-1}(B) = \Psi(B)$.

From this point on we shall assume all the cumulative probability distribution functions with which we deal are left-continuous.

Recall that by λ we mean 1-dimensional Lebesgue measure and that whenever G is a 1-dimensional distribution function, then $\widehat{G}$ stands for the quasi-inverse of G.

We say $C: I^N \to I$ is an N-copula (or just copula for short) if there is a probability measure μ defined on the Borel sets of I^N having the property that $\mu(I^{i-1} \times A \times I^{N-i}) = \lambda(A)$ for all Borel sets of I and all i from 1 to N and related to μ by

$$C(x_1, \cdots, x_N) = \mu([0, x_1) \times \cdots \times [0, x_N)).$$

(Another, more algebraic, characterization of copulas may be found in [1].) We also say that μ is the measure induced on I^N by the copula C.

From this point on till the end of this section we shall assume $F_1, \cdots, F_N$ are given 1-dimensional distribution functions and C is a given n-copula. We define

$$F(x_1, \cdots, x_N) = (F_1(x_1), \cdots, F_N(x_N)),$$
$$\widehat{F}(x_1, \cdots, x_N) = (\widehat{F}_1(x_1), \cdots, \widehat{F}_N(x_N)),$$
$$G(x_1, \cdots, x_N) = C(F_1(x_1), \cdots, F_N(x_N)),$$

μ = the probability measure I^N corresponding to C,

ν = the probability measure induced on $\overline{\mathbf{R}}^N$ by the distribution function G,

$D_i = \{y \in I: F_i$ has the value y over some nondegenerate interval $\}$,

$$D = \bigcup_{i=1}^{N} (I^{i-1} \times D_i \times I^{N-i}).$$

Note that each D_i is countable, hence $\lambda(D_i) = 0$ and thus $\mu(D) \leq \sum_{i=1}^{N} \lambda(D_i) = 0$. Notice also that $\widehat{F}^{-1}$ maps subsets of $\overline{\mathbf{R}}^N$ to subsets of I^N the same as Φ does. As in the 1-dimensional case, these two maps are almost the same in a sense which we now make precise.

PROPOSITION 18. For every $A \subseteq \overline{\mathbf{R}}^N$ we have $\widehat{F}^{-1}(A) \subseteq \Phi(A) \subseteq \widehat{F}^{-1}(A) \cup D$.

We finally attain our main result.

THEOREM; N-DIMENSIONAL CASE. If A is ν-measurable, then $\Phi(A)$ is μ-measurable and $\mu[\Phi(A)] = \nu(A)$. If $B \in \mathfrak{D}_N$ and B is μ-measurable, then $\Psi(B)$ is ν-measurable and $\nu[\Psi(B)] = \mu(B)$.

Proof. Since $\Phi(A)$ differs from $\widehat{F}^{-1}(A)$ by a set of μ-measure zero and $\widehat{F}$ is componentwise nondecreasing, it follows that $\Phi(A)$ is μ-measurable whenever A is ν-measurable. Now let us see how Φ behaves when applied to N-dimensional intervals.

Let

$$A = [-\infty, a_1) \times \cdots \times [-\infty, a_N),$$
$$B = [0, F_1(a_1)) \times \cdots \times [0, F_N(a_N)),$$
and
$$B^+ = [0, F_1(a_1)] \times \cdots \times [0, F_N(a_N)].$$

We will first establish that $B \subseteq \Phi(A) \subseteq B^+$. To do this it is sufficient to show that for a given i we have $[0, F_i(a_i)) \subseteq \Phi_i([-\infty, a_i)) \subseteq [0, F_i(a_i)]$.

Choose $t \in [0, F_i(a_i))$. This means $0 \leq t < F_i(a_i)$. Since F_i is left-continuous, there must be some $x \in [-\infty, a_i)$ for which $F_i(x) \leq t \leq F_i(x^+)$. Thus $t \in \Phi_i([-\infty, a_i))$. Hence $[0, F_i(a_i)) \subseteq \Phi_i([0, a_i))$.

Now choose $t \in \Phi_i([-\infty, a_i))$. There must be some $x \in [-\infty, a_i)$ such that $F_i(x) \leq t \leq F_i(x^+)$. Since F_i is nondecreasing, we have $F_i(x^+) \leq F_i(a_i)$, so that $t \in [0, F_i(a_i)]$. Hence $\Phi_i([-\infty, a_i)) \subseteq [0, F_i(a_i)]$.

We must have $\mu(B) \leq \mu(\Phi(A)) \leq \mu(B^+)$. But B and B^+ differ by sets of μ-measure zero. Therefore we have

$$\begin{aligned}
\mu[\Phi(A)] &= \mu(\Phi([-\infty, a_1) \times \cdots \times [-\infty, a_N)) \\
&= \mu([0, F_1(a_1)) \times \cdots \times [0, F_N(a_N))) \\
&= C(F_1(a_1), \cdots, F_N(a_N)) \\
&= G(a_1, \cdots, a_N) \\
&= \nu([-\infty, a_1) \times \cdots \times [-\infty, a_N)) \\
&= \nu(A).
\end{aligned}$$

Now suppose A has the form $[a_1, b_1) \times \cdots \times [a_N, b_N)$. Using union and complementation, it can be written in terms of intervals of the form $[-\infty, c_1) \times \cdots \times [-\infty, c_N)$ in such a way that $\mu[\Phi(A)] = \nu(A)$ follows easily. See the discussion of n-increasing functions in [1] for details.

The rest of this proof follows as in the 1-dimensional case. □

It is perhaps worthy of note that Φ and Ψ depend only on the marginals $F_1, \cdots, F_N$ and not on the particular μ and ν. Thus any given Φ and Ψ are measure-preserving for a whole family of pairs of probability measures μ and ν.

REFERENCES

1. SCHWEIZER, B., and SKLAR, A. Probabilistic Metric Spaces, North Holland, Amsterdam, 1983.

2. SKLAR, A. Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Statist. Univ. Paris, 8 (1959), 229-231.

3. FRÉCHET, M. Sur les tableaux de corrélation dont les marges sont donnés, Ann. Univ. Lyon 9, Sec. A (1951), 53-77.

4. MIKUSIŃSKI, P., SHERWOOD, H., and TAYLOR, M. Probabilistic interpretations of copulas and their convex sums. Conference Proceedings of the Symposium on Distributions with Given Marginals (Frechet classes), Universita La Sapienza, Rome, Italy, April 4-7, 1990. To be published by Kluwer Academic Publishers.

5. MIKUSIŃSKI, P., SHERWOOD, H., and TAYLOR, M. Shuffles of Min. To appear in Stochastica.

6. MIKUSIŃSKI, P., SHERWOOD, H., and TAYLOR, M. The Fréchet bounds revisited. To appear in Real Anal. Exchange.

7. PHILLIPS, M. On a Transformation which Normalizes Lebesgue-Stieltjes Measures, research report, Dept. of Math., Univ. of Central Florida, 1991.