Abe Sklar

Random variables, joint distribution functions, and copulas

Kybernetika, Vol. 9 (1973), No. 6, (449)--460

Persistent URL: http://dml.cz/dmlcz/125838

KYBERNETIKA — VOLUME 9 (1973), NUMBER 6

Random Variables, Joint Distribution Functions, and Copulas*

ABE SKLAR

If G is an n-dimensional joint distribution function with 1-dimensional margins $F_1, \ldots, F_n$, then there exists a function C (called an "n-copula") from the unit n-cube to the unit interval such that

$$G(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n))$$

for all real n-tuples $(x_1, \ldots, x_n)$. The paper is devoted to an investigation of the structure and properties of n-copulas and their connection with random variables.

1. INTRODUCTION

Multi-dimensional distribution functions, as they appear in probability theory, are complicated objects; and their use has been labelled "an anachronism" ([8], Chapter 1, p. 14) and of "limited value" ([4], v. II, Chapter V, p. 131). Such severe judgments, now widespread, may in time be mollified as the realization spreads that multi-dimensional distribution functions are composite objects built up from simpler things (cf. the remark following Theorem 1). Some of these simpler things, 1-dimensional distribution functions, are well known; but the others, the "copulas" of the title, are not. The major purpose of this paper is to make copulas better known by indicating something of their structure, and actual and potential usefulness.

Special cases of copulas had previously appeared in the literature (e.g. in [5]), but in their general form they were first introduced by the author in [26]. This was in connection with the work of Fréchet and others on joint distribution functions (see, e.g., [7], [3], [5], [18]). The name was chosen to express the fact that copulas embody the way in which multidimensional distribution functions are coupled to their 1-dimensional margins (cf. Theorem 1), and also the way in which random

* Presented at the *Sixth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes,* Prague — September 19—25, 1971.

 variables defined over a common probability space are connected or linked together (cf. Theorems 4–7).

The proofs of the results contained in this paper are all quite elementary; but as some of them are long and cumbersome, limitations of space have forced us to omit them. They will appear elsewhere. Apart from this, every effort has been made to keep the paper self-contained, in particular by defining everything from the beginning. In this connection, one small deviation from current standard practice may be mentioned here: our "1-dimensional distribution functions" are not required to be continuous at $-\infty$ and at $+\infty$; and similarly for multidimensional distribution functions. Nothing is lost in this way, and a certain small measure of generality is gained.

2. DISTRIBUTION FUNCTIONS AND COPULAS

Let n be a positive integer. An *n-interval*

$$N = [(x_1, \ldots, x_n), (y_1, \ldots, y_n)], \quad (x_m \leqq y_m \text{ for all } m \leqq n),$$

is the Cartesian product of the n closed intervals $[x_1, y_1], \ldots, [x_n, y_n]$. In particular, the unit n-cube I^n is the Cartesian n^{th} power of the closed unit interval $I = [0, 1]$, and the extended n-space E^n is the Cartesian n^{th} power of the extended real line $E = [-\infty, +\infty]$. The *vertices* of $[(x_1, \ldots, x_n), (y_1, \ldots, y_n)]$ are the 2^n n-tuples $(z_1, \ldots, z_n)$ such that each z_m is equal to either x_m or y_m. (In a *degenerate* n-interval, some or all of the vertices will coincide.)

An *n-place real function* G is a function whose domain, Dom G, is a non-empty subset of E^n, and whose range, Ran G, is a subset of the unextended real line. Suppose N is an n-interval whose vertices all lie in Dom G. Then the *G-volume* of N is the sum

$$V_G(N) = \sum \alpha(z_1, \ldots, z_n)\, G(z_1, \ldots, z_n) \tag{1}$$

in which the summation is over all 2^n vertices $(z_1, \ldots, z_n)$ of $N = [(x_1, \ldots, x_n), (y_1, \ldots, y_n)]$, and:

$$\alpha(z_1, \ldots, z_n) = \begin{cases} 1 & \text{if } z_m = x_m \text{ for an even number of } m\text{'s}, \\ -1 & \text{if } z_m = x_m \text{ for an odd number of } m\text{'s}. \end{cases} \tag{2}$$

(It follows that $V_G(N) = 0$ if N is degenerate.) We call G *n-increasing* if $V_G(N) \geqq 0$ for all n-intervals N whose vertices lie in Dom G.

Definition 1. An *n-dimensional distribution function* is an n-place real function G whose domain is all of E^n, whose range is a subset of I, and which satisfies the conditions:

$$G(+\infty, \ldots, +\infty) = 1, \tag{3}$$

(4) $$G(x_1, \ldots, x_n) = 0 \quad \text{if} \quad x_m = -\infty \quad \text{for any} \quad m \leqq n\,,$$

(5) G is n-increasing.

For $1 \leqq m \leqq n$, the m^{th} *1-margin* of G is the 1-place real function F_m defined on E by the equation

(6) $$F_m(x_m) = G(+\infty, \ldots, +\infty, x_m, +\infty, \ldots, +\infty)\,.$$

One consequence of (4) and (5) is that an n-dimensional distribution function is non-decreasing in each of its arguments separately. It follows immediately that every 1-margin of an n-dimensional distribution function is itself an ordinary 1-dimensional distribution function.

Definition 2. An *n-dimensional copula* (briefly, an *n-copula*) is an n-place real function C with Dom $C = I^n$, Ran $C = I$, which satisfies the following conditions:

(7) $$C(1, \ldots, 1, x_m, 1, \ldots, 1) = x_m \quad \text{for each} \quad m \leqq n \quad \text{and all} \quad x_m \quad \text{in} \quad I\,,$$

(8) $$C(x_1, \ldots, x_n) = 0 \quad \text{if} \quad x_m = 0 \quad \text{for any} \quad m \leqq n\,,$$

(9) C is n-increasing.

It follows from (7)–(9) that there is a unique 1-copula: the identity function on I; and that every n-copula is non-decreasing in each argument separately, and jointly continuous in all arguments. Furthermore, every n-copula C defines a non-negative Lebesgue-Stieltjes measure λ_C on I^n with $\lambda_C(I^n) = 1$.

Theorem 1. For $n \geqq 2$, let G be an n-dimensional distribution function with 1-margins $F_1, \ldots, F_n$. Then there exists an n-copula C such that

(10) $$G(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n))$$

for all n-tuples $(x_1, \ldots, x_n)$ in E^n. Conversely, let C be an n-copula and $(F_1, \ldots, F_n)$ an n-tuple of 1-dimensional distribution functions. Let an n-place real function G be defined via (10). Then G is an n-dimensional distribution function with 1-margins $F_1, \ldots, F_n$.

Corollary. Let $X_1, \ldots, X_n$ be E-valued random variables defined over a common probability space. Denote the individual distributions functions of $X_1, \ldots, X_n$ by $F_1, \ldots, F_n$, respectively, and the joint distribution function of $X_1, \ldots, X_n$ by G. Then F_m is the m^{th} 1-margin of G for all $m \leqq n$, and there exists an n-copula C (a *connecting copula* of $X_1, \ldots, X_n$) such that (10) holds for all $(x_1, \ldots, x_n)$ in E^n. If each F_m is continuous, then the connecting copula C of $X_1, \ldots, X_n$ is unique.

N.B. In terms of the notions of [21], equation (10) shows that G is the *serial composite* of the n-copula C and a *β-composite* of 1-dimensional distribution functions. Cf. the remarks in the Introduction.

Theorem 1 can be proved via the following sequence of lemmas in each of which n is a fixed integer $\geqq 2$ and G is an n-dimensional distribution function with 1-margins $F_1, \ldots, F_n$:

Lemma 1. Let i, j be integers such that $1 \leqq i < j \leqq n$. Let x_i, y_i, x_j, y_j be elements of E with $x_i \leqq y_i$, $x_j \leqq y_j$, and $z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_{j-1}, z_{j+1}, \ldots, z_n$ be arbitrary elements of E. Then we have:

$$(11) \qquad G(z_1, \ldots, x_i, \ldots, x_j, \ldots, z_n) - G(z_1, \ldots, x_i, \ldots, y_j, \ldots, z_n) - G(z_1, \ldots, y_i, \ldots, x_j, \ldots, z_n) + G(z_1, \ldots, y_i, \ldots, y_j, \ldots, z_n) \geqq 0 .$$

Lemma 2. Let m be a positive integer $\leqq n$, and $x_m, y_m, z_1, \ldots, z_{m-1}, z_{m+1}, \ldots, z_n$ arbitrary elements of E. Then we have:

$$(12) \qquad |G(z_1, \ldots, z_{m-1}, x_m, z_{m+1}, \ldots, z_n) - G(z_1, \ldots, z_{m-1}, y_m, z_{m+1}, \ldots, z_n)| \leqq \leqq |F_m(x_m) - F_m(y_m)| .$$

Lemma 3. Let $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$ be arbitrary n-tuples in E^n. Then we have:

$$(13) \qquad |G(x_1, \ldots, x_n) - G(y_1, \ldots, y_n)| \leqq |F_1(x_1) - F_1(y_1)| + \ldots \ldots + |F_n(x_n) - F_n(y_n)| .$$

Lemma 4. The set of pairs

$$\{((F_1(x_1), \ldots, F_n(x_n)), G(x_1, \ldots, x_n)| (x_1, \ldots, x_n) \text{ in } E^n\}$$

is the graph of an n-place real function C^* such that:

$$\operatorname{Dom} C^* = \operatorname{Ran} F_1 \times \ldots \times \operatorname{Ran} F_n , \quad \operatorname{Ran} C^* = \operatorname{Ran} G \subseteq I .$$

Hence we have

$$(14) \qquad G(x_1, \ldots, x_n) = C^*(F_1(x_1), \ldots, F_n(x_n)) \quad \text{for all} \quad (x_1, \ldots, x_n) \quad \text{in} \quad I^n .$$

Lemma 5. The function C^* of Lemma 4 can be extended (generally in many ways) to an n-copula C. If $\operatorname{Ran} F_m = I$ for all $m \leqq n$ (i.e., if each F_m is continuous), then $C^* = C$ and C is unique.

From Theorem 1, by applying Kolmogorov's fundamental consistency theorem ([14], Ch. III, § 4) we derive:

Theorem 2. Let C be an n-copula and $(F_1, \ldots, F_n)$ an n-tuple of 1-dimensional distribution functions. Then there exists a probability space and E-valued random variables $X_1, \ldots, X_n$ defined over that space such that C is a connecting copula of $X_1, \ldots, X_n$, and F_m is the distribution function of X_m for each $m \leqq n$.

Definition 3. Let T be a 2-place real function. The *serial iterates* of T (cf. [21]) are the functions T^n defined by the following recursion scheme:

$$T^1 = T; \tag{15}$$

$$T^n(x_1, \ldots, x_{n+1}) = T(T^{n-1}(x_1, \ldots, x_n), x_{n+1}) \quad \text{for} \quad n \geqq 2 .$$

Thus, for each $n \geqq 2$, T^{n-1} is an n-place real function.

We let Prod and Min, respectively, denote the restrictions of the ordinary real product and minimum functions to I^2, so that:

$$\text{Prod}(x, y) = x \cdot y , \tag{16}$$

$$\text{Min}(x, y) = \begin{cases} x , & x \leqq y , \\ y , & x \geqq y , \end{cases} \tag{17}$$

for all (x, y) in I^2. Correspondingly, we define a function Tm on I^2 by the equation:

$$\text{Tm}(x, y) = \max(x + y - 1, 0) \quad \text{for all} \quad (x, y) \quad \text{in} \quad I^2 . \tag{18}$$

Theorem 3. Let n be an integer $\geqq 2$. Then Prod^{n-1} and Min^{n-1} are each n-copulas, while Tm^{n-1} is an n-copula if and only if $n = 2$. Any n-copula C satisfies the inequalities

$$\text{Tm}^{n-1} \leqq C \leqq \text{Min}^{n-1} ; \tag{19}$$

and the left-hand inequality in (19) is best-possible in the sense that for any n-tuple $(x_1, \ldots, x_n)$ in I^2, there is an n-copula C_1 such that $C_1(x_1, \ldots, x_n) = \text{Tm}^{n-1}(x_1, \ldots, x_n)$.

3. COPULAS AND RANDOM VARIABLES

An immediate consequence of the definition of independence is:

Theorem 4. Let n be an integer $\geqq 2$, and $X_1, \ldots, X_n$ be E-valued random variables with *continuous* distribution functions. Then $X_1, \ldots, X_n$ are independent if and only if their (unique) connecting copula is Prod^{n-1}.

A related but somewhat less immediate result is contained in:

Theorem 5. Let n and $X_1, \ldots, X_n$ be as in Theorem 4. Then every $(n-1)$-element subset of $\{X_1, \ldots, X_n\}$ is a set of independent random variables if and only if the connecting copula C of $X_1, \ldots, X_n$ satisfies the inequalities:

$$(20) \quad \begin{aligned} &\text{Prod}^{n-1}(x_1, \ldots, x_n) - \text{Prod}^{n-1}(1-x_1, \ldots, 1-x_n) \leqq C(x_1, \ldots, x_n) \leqq \\ &\leqq \text{Prod}^{n-1}(x_1, \ldots, x_n) + \min\big(x_1(1-x_2)\ldots(1-x_n), \ldots, (1-x_1)\ldots \\ &\ldots(1-x_{n-1})\,x_n\big), \quad (\text{for } n \text{ even}); \end{aligned}$$

$$(21) \quad \begin{aligned} &\text{Prod}^{n-1}(x_1, \ldots, x_n) - \min\big(x_1(1-x_2)\ldots(1-x_n), \ldots, (1-x_1)\ldots \\ &\ldots(1-x_{n-1})\,x_n\big) \leqq C(x_1, \ldots, x_n) \leqq \text{Prod}^{n-1}(x_1, \ldots, x_n) + \\ &+ \text{Prod}^{n-1}(1-x_1, \ldots, 1-x_n), \quad (\text{for } n \text{ odd}). \end{aligned}$$

Theorem 6. Let n and $X_1, \ldots, X_n$ be as in Theorem 4. Then each of the random variables $X_1, \ldots, X_n$ is an *increasing* function a.e. of any of the others if and only if their connecting copula is Min^{n-1}. In particular, if X_1 and X_2 are random variables with respective distribution functions F_1 and F_2 that are both continuous, then $X_1 = X_2$ a.e. if and only if $F_1 = F_2$ and the connecting copula of X_1 and X_2 is Min.

Corollary. If $X_1, \ldots, X_n$ are all strictly increasing, or all strictly decreasing, functions on the unit interval endowed with Lebesgue measure, then the unique connecting copula of $X_1, \ldots, X_n$ is Min^{n-1}.

Theorem 7. Let X_1, X_2 be E-valued random variables with respective distribution functions F_1, F_2 that are continuous. Then X_1 and X_2 are *decreasing* functions of each other a.e. if and only if the connecting copula of X_1 and X_2 is Tm. In particular, $X_1 + X_2 = c$ (a.e), where c is some real constant, if and only if $F_1(x) + F_2(c-x) = 1$ for all x in E, and the connecting copula of X_1 and X_2 is Tm.

Corollary. If X_1 is a strictly increasing, and X_2 a strictly decreasing function on the unit interval endowed with Lebesgue measure, then the unique connecting copula of X_1 and X_2 is Tm.

4. ASSOCIATIVE COPULAS

The preceding results indicate the desirability of obtaining structural characterizations of copulas. To date this has been done only for *associative* 2-copulas and serial iterates of certain associative 2-copulas. Here the meaning of "associativity" is the usual one: a 2-copula C is associative if it satisfies the condition:

$$(22) \quad C(C(x_1, x_2), x_3) = C(x_1, C(x_2, x_3)) \quad \text{for all} \quad (x_1, x_2, x_3) \quad \text{in} \quad I^3 .$$

In providing the characterization of associative 2-copulas we begin with

Definition 4. Let g be an E-valued function defined, continuous, and strictly decreasing on I, with $g(1) = 0$. Let f be the function defined on the half-line $[0, +\infty]$ by the requirement that f coincides with g^{-1} on $[0, g(0)]$, while $f(x) = 0$ for x in $[g(0), +\infty]$. (Note that each of the functions f, g determines the other.) Let T be a 2-place real function with Dom $T = I^2$. Then f is an *outer generator*, and g an *inner generator* of T, if T admits the representation

$$(23) \qquad T(x, y) = f(g(x) + g(y)) \quad \text{for all} \quad (x, y) \quad \text{in} \quad I^2\,;$$

and if T admits the representation (23), then T itself is called *Archimedean* (cf. [15]).

N.B. If g is an inner generator of T, then any positive multiple of g is also an inner generator of T; conversely, any 2 inner generators of the same Archimedean function are positive multiples of one another.

Definition 5. Let S be a collection of 2-place real functions, each defined on I^2, and each assuming the value 0 at the argument $(0, 0)$ and the value 1 at the argument $(1, 1)$. Let T be a 2-place real function with Dom $T = I^2$. Then T is an *ordinal sum* of elements of S if there exists a family $\{(a_i, b_i)\}$ of pairwise disjoint open subintervals of I such that for each interval (a_i, b_i) in the family there is a unique function T_i in S, and:

$$(24) \qquad T(x, y) = a_i + (b_i - a_i)\, T_i\left(\frac{x - a_i}{b_i - a_i}, \frac{y - a_i}{b_i - a_i}\right)$$

for all (x, y) in the (closed) 2-interval $[a_i, b_i]^2$; while

$$(25) \qquad T(x, y) = \mathrm{Min}\,(x, y)$$

if (x, y) lies in none of the 2-intervals $[a_i, b_i]^2$.

We now combine some results of C.-H. Ling in [15], and B. Schweizer and the author in [20] to obtain:

Theorem 8. Let C be an associative 2-copula. Then:
(a) $C(x, x) = x$ for all x in I if and only if $C = \mathrm{Min}$;
(b) $C(x, x) < x$ for all x in the open interval $(0, 1)$ if and only if C is Archimedean and every outer generator of C is convex;
(c) $C(x, x) < x$ for some, but not all x in $(0, 1)$ if and only if C is an ordinal sum of Archimedean 2-copulas, i.e., Archimedean functions with convex outer generators.

Theorem 9. Let C be an Archimedean 2-copula. Then the serial iterates C^{n-1} of C are n-copulas for all $n \geqq 2$ if and only if there is an outer generator f of C that is *completely monotonic*, i.e., if and only if f is an analytic function satisfying the conditions:

$$(-1)^n f^{(n)} \geqq 0 \quad \text{for all} \quad n \geqq 0\,, \tag{26}$$

(cf. [30], Chapter IV, and [4], v. II, Chapter XIII, § 4). If one outer generator of C is completely monotonic, then all are, and C satisfies the inequalities Prod $\leqq C <$ $<$ Min.

Theorem 9 has been extended and applied to sets of exchangeable random variables by C. H. Kimberling [13].

5. OPERATIONS ON SPACES OF DISTRIBUTION FUNCTIONS

In working simultaneously with many 1-dimensional distribution functions, it is usually convenient to normalize them in some way. We normalize 1-dimensional distribution functions by requiring them to be right-continuous on the *unextended* real line; we then call the set of all such normalized functions Δ. The important subset of Δ consisting of all normalized F such that $F(0^-) = 0$ will be denoted by Δ^+. Under suitable topologies, in particular that of weak convergence (see [8], Chapter 2, § 9), Δ and Δ^+ become topological spaces.

Binary operations on Δ and Δ^+ are natural objects of study, and many such operations are definable in terms of 2-copulas. The most direct way is to let each 2-copula C induce an operation π_C on Δ (resp., Δ^+) via:

$$(\pi_C(F_1, F_2))(x) = C(F_1(x), F_2(x)) \tag{27}$$

for all F_1, F_2 in Δ (resp., Δ^+) and all x in E.

Operations on random variables induce corresponding families of operations on Δ and Δ^+, as follows: Let X_1, X_2 be random variables with respective distribution functions F_1, F_2 and connecting copula C. If H is a measurable 2-place real function, then $H(X_1, X_2)$ is a random variable whose distribution function depends not only on H, F_1, and F_2, but on C as well. If we denote this distribution function by $\langle H\rangle_C (F_1, F_2)$, then its value for a real argument x is given by the expression:

$$(\langle H\rangle_C (F_1, F_2))(x) = \iint_{H(u,v) \leqq x} \mathrm{d}C(F_1(u), F_2(v))\,. \tag{28}$$

Thus H and C together yield a binary operation $\langle H\rangle_C$ on Δ. If H is non-negative on the first quadrant, then the restriction of $\langle H\rangle_C$ to Δ^+ is a binary operation on Δ^+.

Particular interest attaches to the case when H is addition, i.e., when $H(u, v) =$ $= u + v$. In this case, we shall write σ_C instead of $\langle H\rangle_C$. From (28) it is readily seen that σ_{Prod} is *convolution*. Apart from this supremely important and intensively studied case (for a comprehensive treatment of which, see [16]), very little is known about σ_C. The most recent general result is the following remarkable theorem due to M. J. Frank [6]:

Theorem 10. The operation σ_C is associative if and only if: either $C = \text{Min}$, or $C = \text{Prod}$, or C is an ordinal sum of Prod's alone.

A key step in Frank's proof is the establishment of the following:

Lemma 6. A necessary (though definitely not sufficient) condition for σ_C to be associative is that both the functions C and $\bar{C}$ be associative, where $\bar{C}$ is the function defined in terms of C by:

(29) $$\bar{C}(x, y) = x + y - C(x, y).$$

The operations π_C are included among the operations $\langle H \rangle_C$ by virtue of the easily verified identity

(30) $$\pi_C = \langle \max \rangle_C .$$

There are, however, operations on Δ and Δ^+ that *cannot* be obtained as special cases of $\langle H \rangle_C$. The most notable of these are the operations τ_C introduced by A. N. Šerstnev in the course of his work on probabilistic normed spaces and probabilistic metric spaces (see, e.g., [24], [25]). These are defined for each fixed 2-copula C via the equations:

(31) $$(\tau_C(F_1, F_2))(x) = \sup \{C(F_1(u), F_2(v)) \mid u + v = x\},$$

for any F_1, F_2 in Δ (or Δ^+) and all real x.

Theorem 11. The operation τ_{Min} coincides with the operation σ_{Min}. But if $C \neq \text{Min}$, then the operation τ_C does *not* coincide with the operation $\langle H \rangle_C$ for *any* measurable 2-place function H.

Since the operations τ_C are commutative, associative, monotonic, continuous in the topology of weak convergence, etc., this shows that there are very natural operations on distribution functions that do not correspond in any simple fashion to operations on random variables. On the other hand, the operations τ_C *are* related to the other operations via *inequalities*, as follows:

Theorem 12. Let C be a 2-copula. Then, on Δ^+, we have:

(32) $$\tau_{\text{Tm}} \leqq \tau_C \leqq \sigma_C \leqq \pi_C \leqq \pi_{\text{Min}},$$

in the sense that $\tau_C(F_1, F_2) \leqq \sigma_C(F_1, F_2)$ for all F_1, F_2 in Δ^+, etc. All the inequalities in (32), with the exception of $\sigma_C \leqq \pi_C$, remain valid in Δ.

In 1942, K. Menger initiated the theory of *probabilistic metric spaces* ([17]; for a survey, see [19]). In all versions of the theory, the basic notion is a function that maps pairs of points (p, q) in an abstract space S into distribution functions F_{pq} in Δ^+; the value $F_{pq}(x)$ of F_{pq} for a real argument x is then interpreted as the probability that the distance between p and q is $\leqq x$. To sustain this interpretation, some form of triangle inequality among the F_{pq}'s is assumed.

In 1956, A. Špaček introduced the notion of a *random metric space* [27]. Although he continued to write on the subject (see [28]) his untimely death interrupted its subsequent development for many years, and it is only now that active work in the area is resuming ([1], [2]).

Let S be a non-empty set, and $(\Omega, \mathscr{B}, P)$ a probability space. An *Ω-random metric* on S is a function ξ from $S^2 \times \Omega$ into E such that:

(a) $\xi(\cdot, \omega) = d_\omega$ is an (ordinary) metric on S for almost all ω in Ω;
(b) $\xi((p, q), \cdot) = X_{pq}$ is a random variable over $(\Omega, \mathscr{B}, P)$ for all pairs (p, q) in S^2.

It follows that if (p, q, r) is any triple in S^3, then we have:

$$X_{pr} \leqq X_{pq} + X_{qr} \quad \text{a.e.} \tag{33}$$

Now if we go from the random variables X_{pq} to their distribution functions F_{pq}, we obtain a probabilistic metric space corresponding to the original random metric. We seek the form the triangle inequality takes in this associated probabilistic metric space. To this end, let C_{pqr} be a connecting copula of the random variables X_{pq} and X_{qr}. Then (28) and (33) yield

$$F_{pr} \geqq \sigma_{C_{pqr}}(F_{pq}, F_{qr})\,; \tag{34}$$

and from this, an appeal to (32) yields

$$F_{pr} \geqq \tau_{\mathrm{Tm}}(F_{pq}, F_{qr})\,. \tag{35}$$

Thus (35) will hold in any probabilistic metric space derived from a random metric; and an adaptation of the methods of E. Thorp in [29] can be made to show that (35) is the best-possible universal form of the triangle inequality valid in such spaces. In the other direction, and essentially as an application of Theorem 11, it can be shown that there are probabilistic metric spaces in which (35) holds, but that *cannot* be derived from random metrics.

Similar results have recently been obtained in the context of information theory. In the past few years, J. Kampé de Fériet and B. Forte have begun to construct an axiomatic theory of information independent of probability theory ([11]; see also [9], [12]). Motivated by the analogy with metric spaces, B. Schweizer and the author introduced the notion of a *probabilistic information space* [22]. Kampé de Fériet then introduced *random information spaces*, with the very striking interpretation of the probability space involved as an "ensemble d'observateurs" [10]. In [23], B. Schweizer and the author showed that the distribution functions in probabilistic information spaces derived from random information spaces satisfy certain inequalities that are best-possible universal inequalities for such spaces; and that there are probabilistic information spaces in which these inequalities hold that *cannot* be derived from random information spaces.

(Received February 1, 1972.)

REFERENCES

[1] J. B. Brown: Random metrics for statistical metric spaces. Z. Wahrsch. *24* (1972), 49—62.

[2] P. Calabrese: A random variable approach to probabilistic metric spaces. To appear.

[3] G. Dall'Aglio: Sulla compatibilità delle funzioni di ripartizione doppia. Rendiconti Mat. *18* (1959), 385—413.

[4] W. Feller: An introduction to probability theory and its applications. Wiley, New York 1966.

[5] R. Féron: Sur les tableaux de corrélation dont les marges sont données. Publ. Inst. Statistique Univ. Paris *5* (1956), 1—12.

[6] M. J. Frank: Associativity of operators on spaces of distribution functions. Notices Amer. Math. Soc. *19* (1972), 1 A-143, Abstract 691—39—5. To appear in Aequat. Math.

[7] M. Fréchet: Sur les tableaux de corrélation dont les marges sont données. Ann. Univ. Lyon *9* (1951), Sec. A, 53—77.

[8] B. V. Gnedenko and A. N. Kolmogorov: Limit distributions for sums of independent random variables. Addison-Wesley, Reading 1954.

[9] J. Kampé de Fériet: Mesure de l'information fournie par un événement. In: Colloques Internat. C.N.R.S. No. 186, Clermont-Ferrand 1969. C.N.R.S., Paris 1970, 191—221.

[10] J. Kampé de Fériet: Mesures de l'information par un ensemble d'observateurs. C. R. Acad. Sci. Paris *269* (1969), Sér. A, 1081—1085; ibid. *271* (1971), Sér. A, 1017—1021.

[11] J. Kampé de Fériet et B. Forte: Information et probabilité. C. R. Acad. Sci. Paris *265* (1967), Sér. A, 110—114, 142—146, 350—353.

[12] J. Kampé de Fériet, B. Forte et P. Benvenuti: Forme générale de l'opération de composition continue d'une information. C. R. Acad. Sci. Paris *269* (1969), Sér. A, 529—534.

[13] C. H. Kimberling: A probabilistic interpretation of complete monotonicity. To appear.

[14] A. N. Kolmogorov: Foundations of the theory of probability. Second English edition. Chelsea, New York 1956.

[15] C.-H. Ling: Representation of associative functions. Publicat. Math. Debrecen *12* (1965), 189—212.

[16] Y. V. Linnik: Décompositions des lois de probabilités. Gauthier-Villars, Paris 1962.

[17] K. Menger: Statistical metrics. Proc. Nat. Acad. Sci. U.S.A. *28* (1942), 535—537.

[18] A. Nataf: Détermination des distributions de probabilités dont les marges sont données. C. R. Acad. Sci. Paris *255* (1962), 42—43.

[19] B. Schweizer: Probabilistic metric spaces — the first 25 years. New York Statistician *19* (1967), 2, 3—6.

[20] B. Schweizer and A. Sklar: Associative functions and statistical triangle inequalities. Publicat. Math. Debrecen *8* (1961), 169—186.

[21] B. Schweizer and A. Sklar: A grammar of functions. Aequat. Math. *2* (1968), 62—85; ibid. *3* (1969), 15—43.

[22] B. Schweizer and A. Sklar: Mesures aléatoires de l'information. C. R. Acad. Sci. Paris *269* (1969), Sér. A, 721—723.

[23] B. Schweizer and A. Sklar: Mesure aléatoire de l'information et mesure de l'information par un ensemble d'observateurs. C. R. Acad. Sci. Paris *272* (1971), Sér. A, 149—152.

[24] A. N. Šerstnev: On the notion of a random normed space. Doklady Akad. Nauk S.S.S.R. *149* (1963), 280—283; translated in: Soviet Math. Doklady *4* (1963), 388—391.

[25] A. N. Šerstnev: On probabilistic generalizations of metric spaces. Učen. Zap. Kazan. Gos. Univ. *124* (1964), 2, 3—11.

[26] A. Sklar: Fonctions de répartition à *n* dimensions et leurs marges. Publ. Inst. Statistique Univ. Paris *8* (1959), 229—231.

[27] A. Špaček: Note on K. Menger's probabilistic geometry. Czechoslovak Math. J. *6* (1956), 72—74.

[28] A. Špaček: Random metric spaces. In: Transactions of the Second Prague Conference, Prague 1959. NČSAV, Praha and Academic Press, New York 1960, 627—628.

[29] E. Thorp: Best-possible triangle inequalities for statistical metric spaces. Proc. Amer. Math. Soc. *11* (1960), 734—740.

[30] D. V. Widder: The Laplace transform. Princeton Univ. Press, Princeton 1941.

Abe Sklar, Professor of Mathematics, Illinois Institute of Technology, Chicago, Illinois 60616. U.S.A.