STATISTICAL METRIC SPACES

by

GARY ALLAN BOGAR

B. S., Missouri State Teachers College, 1960

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Mathematics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1963

Approved by:

William L. Stamey
Major Professor

TABLE OF CONTENTS

INTRODUCTION.. 1

DEFINITIONS AND PROPERTIES.. 2

Statistical Metric Space.. 2

Menger Space.. 5

Wald Space...15

Equilateral Spaces...18

Simple Spaces..20

TOPOLOGY, CONVERGENCE, CONTINUITY...................................25

Neighborhood...25

Hausdorff Space..26

Uniformity...33

INTRODUCTION

Abstract metric spaces were introduced by M. Frechet in 1906. The study of such spaces furnishes a common idealization of a large number of mathematical, physical, and other scientific constructs in which the idea of "distance" appears. The objects under consideration in such a space may be quite arbitrary, but are generally called "points" for suggestiveness. The objective is to associate a non-negative real number with each pair of points of the given set in such a way that the numbers associate with pairs of triples of points satisfy certain conditions. The association of a single real non-negative number with the pairs is an over idealized situation. In practice the distance between two points is not given by a single measurement but instead is taken as the average of the distances obtained from several measurements. Thus, in a metric space, the distance concept defined by Frechet may be considered as statistical instead of determinate. In place of associating a single number $d(p,q)$, which is read the distance from p to q, with every pair of elements, one may associate a distribution function F_{pq}. For every positive number x, interpret $F_{pq}(x)$ as the probability that the distance from p to q is less than x. Upon associating a distribution function with every pair of elements p and q, a generalization of the concept of a metric space is obtained. This generalization was introduced by K. Menger and is referred to as a statistical metric space.

In the original paper K. Menger [1] presented four postulates, using distribution functions, which are comparable to the four postulates for a metric space given by Frechet. At this time K. Menger also discussed the theory of betweenness in a space S and possible fields of application.

Shortly after Menger's paper, A. Wald published a paper in which he criticized Menger's generalized triangular inequality and proposed a substitute. Using this new postulate, Wald was able to construct a theory of betweenness having certain advantages over Menger's theory.

The next extensive work published on the subject was by B. Schweizer and A. Sklar [3]. In this paper they examined the spaces defined by Menger and Wald and presented necessary conditions for the equivalence of the two spaces. The remainder of the paper was devoted to the study of specific spaces and the topological properties of a statistical metric space.

In 1961, E. Thorp published a paper in which he discussed the properties of the t-functions defined by Menger. At this time Thorp was able to prove that given a t-function there is a space for which it is the strongest. He also worked on the problem, given a statistical metric space can a strongest t-function be found?

This paper will be divided into the following three sections.

(I) The study of the axioms for a statistical metric space, with particular emphasis on the triangle inequality.

(II) Study of particular spaces.

(III) Topological properties of a statistical metric space and the continuity properties of the distance distribution function.

DEFINITIONS AND PROPERTIES

1. Statistical Metric Space. The distributions used here will have the usual properties. A distribution function is any real valued function which is defined on the real line, is left continuous, is non-decreasing, and has greatest lower bound zero and least upper bound one. One particular distribution function will be of interest throughout the paper. It is the function H defined by

$$H(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0. \end{cases}$$

A distribution function F will adhere to the convention that for any x,

$$\lim_{x \to +\infty} F(x) = F(+\infty) = 1$$

and

$$\lim_{x \to -\infty} F(x) = F(-\infty) = 0.$$

In order to compare the definition of a statistical metric space with the definition of a metric space, the postulates due to Frechet are listed here. A metric space (M-space) is an abstract set S of elements called "points" and a mapping d of $S \times S$ into the non-negative real numbers. The mapping d has the following properties:

i. If p and q are elements of S, then $d(p,q) = 0$ if and only if $p = q$; (Identity)

ii. If p and q are elements of S and $p \neq q$, then $d(p,q) > 0$; (Positive)

iii. If p and q are elements of S, then $d(p,q) = d(q,p)$; (Symmetry)

iv. If p, q, and r are elements of S, then $d(p,r) \leq d(p,q) + d(q,r)$. (Triangle inequality)

Definition 1. A statistical metric space (an SM-space) is an ordered pair $(S,\mathcal{F})$ where S is an abstract set (whose elements will be called points) and $\mathcal{F}$ is a mapping of $S \times S$ into a set of distribution functions; therefore, $\mathcal{F}$ associates a distribution function $\mathcal{F}(p,q)$ with every pair (p,q) of points in S. Denote the distribution function $\mathcal{F}(p,q)$ by F_{pq}. Hence $F_{pq}(x)$ denotes the value of F_{pq} for the real argument x. The functions F_{pq} are assumed to satisfy the following conditions:

I. $F_{pq}(x) = 1$ for all $x > 0$ if and only if $p = q$.

II. $F_{pq}(0) = 0$

III. $F_{pq} = F_{qp}$

IV. If $F_{pq}(x) = 1$ and $F_{qr}(y) = 1$, then $F_{pr}(x + y) = 1$.

The following example shows that every metric space may be regarded as an SM-space of a special kind.

Example 1. Set $F_{pq}(x) = H(x - d(p,q))$ for every pair of points (p,q) in an M-space.

1. $F_{pq}(x) = H(x - d(p,q)) = H(x) = 1$ for $x > 0$ and $p = q$, hence $F_{pq}(x) = 1$ for all $x \quad 0$ and $p = q$.

2. $F_{pq}(0) = H(-d(p,q)) = 0$, since $d(p,q) \geq 0$; hence $F_{pq}(0) = 0$.

3. $F_{pq}(x) = H(x - d(p,q)) = H(x - d(q,p)) = F_{qp}(x)$; hence $F_{pq}(x) = F_{qp}(x)$.

4. If $F_{qr}(y) = 1$, then $H(y - d(q,r)) = 1$; therefore, $y > d(q,r)$.

   If $F_{pq}(x) = 1$, then $H(x - d(p,q)) = 1$; therefore, $x > d(p,q)$.

   $F_{pr}(x + y) = H(x + y - d(p,r))$,

   $d(p,q) + d(q,r) > d(p,r)$,

   $x + y > d(p,r)$,

   $H(x + y - d(p,r)) = 1$;

   hence $F_{pr}(x + y) = 1$.

From the interpretation of $F_{pq}(x)$ as the probability that the distance between p and q is less than x, one can see that conditions I, II, and III are generalizations of conditions i, ii, and iii; while IV is the minimal generalization of iv.

In the SM-spaces in which the equality $F_{pq}(x) = 1$ does not hold for any finite x, IV is satisfied only by the null set. It is, therefore, of interest to have a stronger generalization of the triangular inequality. Before discussing two stronger substitutes for postulate IV, it is convenient to make the following definition.

Definition 2. A triangle inequality is said to hold universally in an SM-space if and only if it holds for all triples of points, distinct or not, in that space.

2. Menger Spaces. Menger gave as a generalized triangle inequality the following:

IV_m. $F_{pr}(x + y) \geq T(F_{pq}(x), F_{qr}(y))$ for all x,y,

where T is a 2-place function on the unit square satisfying:

(a) $0 \leq T(a,b) \leq 1$

(b) $T(c,d) \geq T(a,b)$ for $c \geq a$, $d \geq b$

(c) $T(a,b) = T(b,a)$

(d) $T(1,1) = 1$

(e) $T(a,1) > 0$ for $a > 0$.

Due to the rather generalized nature of T, not too much can be said about this function. The interpretation of T is made more precise by choosing a specific function. Some simple examples follow:

T_1: $T(a,b) = \text{Max}(a + b - 1,0)$	i.e., T = Max(Sum - 1,0)
T_2: $T(a,b) = ab$	i.e., T = Product
T_3: $T(a,b) = \text{Min}(a,b)$	i.e., T = Min
T_4: $T(a,b) = \text{Max}(a,b)$	i.e., T = Max
T_5: $T(a,b) = a + b - ab$	i.e., T = Sum - Product
T_6: $T(a,b) = \text{Min}(a + b,1)$	i.e., T = Min(Sum,1)

The six functions are listed in order of increasing strength, where T_m is stronger than T_n if $T_m(a,b) \geq T_n(a,b)$ for all (a,b) on the unit square with strict inequality for at least one pair (a,b). If IV_m holds for any given T, then it will hold for all weaker T's. For T = Min (Max), the interpretation is: The probability that the distance from p to r is less than x + y is not less than the smaller (larger) of the probability that the distance from p to q is less than x and distance from q to r is less than y. Similar interpretation can be given to the other choices of T. Nevertheless, as the following lemmas indicate, the three functions T_4, T_5, and T_6 are actually too strong for most purposes.

Lemma 1. If an SM-space contains two distinct points, then IV_m cannot hold universally in the space under the choice T = Max.

Proof: Let p and q be elements of S such that $p \neq q$. Also let x and y be real arguments satisfying $0 < y < x$.

Suppose IV_m is true for every value of x and y with T = Max. Then $F_{pq}(x) \geq \text{Max}(F_{pq}(x - y), F_{qq}(y)) = 1$. Since x can be any positive number, then by postulate I, p = q. This is a contradiction of the assumption $p \neq q$; hence, it follows that IV_m cannot hold universally in the space under the choice T = Max.

Lemma 2. If an SM-space is not an M-space and if IV_m holds universally in the space for some choice of T satisfying conditions (a)-(e), then the function T has the property that there exists a number a, $0 < a < 1$, such that $T(a,1) > a$.

Proof: If an SM-space is not an M-space, then there is at least one pair p,q of distinct points, for which F_{pq} has a value other than zero or one. F_{pq} being left continuous and monotonic implies that there is an open interval (x,y) on which $0 < F_{pq} \leq 1$. Assume that for $0 < a < 1$, $T(a,1) = a + g(a)$ where $g(a) \geq 0$. Let z be any point in (x,y) and take $t > 0$. Then

$$F_{pq}(z + t) \geq T(F_{pq}(z), F_{pp}(t))$$

$$\geq T(F_{pq}, 1)$$

$$\geq F_{pq}(z) + g(F_{pq}(z)).$$

Letting t approach zero through the positive values gives

$$F_{pq}(z+) \geq F_{pq}(z) + g(F_{pq}(z)) \geq F_{pq}(z).$$

Thus, F_{pq} is discontinuous at a; and therefore, at every point of the open interval (x,y). This is a contradiction, since a non-decreasing function can be discontinuous at only denumerably many points.

Lemma 3. If IV_m holds universally in an SM-space and if T is continuous, then for any $x > 0$, $T(F_{pq}(x),1) \leq F_{pq}(x)$.

Proof: Let points p,q in S and a positive number x be given; choose y such that $0 \leq y < x$.

$$F_{pq}(x) \geq T(F_{pq}(x - y), F_{pp}(y)) = T(F_{pq}(x - y),1).$$

Letting y approach zero from above gives

$$F_{pq}(x) \geq \lim_{y \to 0+} T(F_{pq}(x - y),1).$$

By the assumed continuity of T,

$$F_{pq}(x) \geq \lim_{y \to 0+} T(F_{pq}(x - y),1) = T(\lim_{y \to 0+} F_{pq}(x - y),1)$$

while by the left continuity of F_{pq},

$$F_{pq}(x) \geq T(\lim_{y \to 0+} F_{pq}(x - y),1) = T(F_{pq}(x),1).$$

Hence $F_{pq}(x) \geq T(F_{pq}(x),1)$.

By Lemmas 1, 2, and 3, it follows that T_1, T_2, and T_3 all satisfy the condition $T(a,1) = a$; therefore, properties a, d, and e can be replaced by

$$(a') \quad T(a,1) = a, \ T(0,0) = 0.$$

Since $0 < a \leq 1$ and $0 < b \leq 1$, it follows that

$$T(a,b) \leq T(a,1) = a$$

$$T(a,b) = T(b,a) \leq T(b,1) = b;$$

hence $T(a,b) \leq \text{Min}(a,b)$. Using condition (a') $T = \text{Min}$ becomes the strongest universal T. Similarly the weakest possible T satisfying (a'), (b), and (c) is the function T_W, which is given by:

$$T_W(x,y) = \begin{cases} a & x = a,\ y = 1 \text{ or } y = a,\ x = 1 \\ 0 & \text{otherwise.} \end{cases}$$

One more property is now added to (a'), (b), and (c). This is the associativity condition,

$$\text{(d')} \quad T\left[T(a,b),c\right] = T\left[a,T(b,c)\right]$$

which allows the extension of IV_m to a polygonal inequality.

Definition 3. A Menger space is an SM-space in which IV_m holds universally for some choice of T satisfying conditions (a'), (b), (c), and (d').

The following lemma shows that only distinct triples need be considered in determining whether an SM-space is a Menger space.

Lemma 4. If the points p, q, and r are not all distinct, then IV_m holds for all triples p, q, and r under the choice T satisfying (a'), (b), (c), and (d').

Proof: Choose $T = \text{Min}$.

Case I. If $p = r \neq q$, then $F_{pr}(x + y) = H$; hence, $T = \text{Min}$. implies that $F_{pr}(x + y) \geq \text{Min}\left[F_{pq}(x),\ F_{pr}(y)\right]$.

Case II. If $p = q \neq r$, then for $x,y \geq 0$

$$\text{Min}(F_{pq}(x),\ F_{qr}(y)) = \text{Min}(H(x),F_{qr}(y))$$
$$\leq F_{qr}(y) \leq F_{qr}(x + y) = F_{pr}(x + y).$$

A real-valued function T that satisfies conditions (a'), (b), (c), and (d'), and whose domain is the set of real number pairs (x,y), such that $0 \leq x,y \leq 1$, is called a t-function.

Definition 4. The t-function T' is stronger than the t-function T'', denoted by $T' \geq T''$, if $T'(x,y) \geq T''(x,y)$ for $0 \leq x,y \leq 1$; T' is strictly stronger than T'' if T' is stronger than T'' and there is at least one pair of numbers (x,y) such that $T'(x,y) > T''(x,y)$. Correspondingly, T'' is a weaker or strictly weaker than T'.

In general the stronger the t-function, the more information one has about the SM-space. It is natural to ask the following question. If given an SM-space that can be made into a Menger space, is there a strongest t-function? It is also interesting to determine whether, for a specific t-function there is a corresponding space for which it is the strongest.

The latter problem can be completely solved by constructing, for a given t-function, a space such that it is strongest. The general case for which it is proved will contain an uncountable number of points. Nevertheless, for some t-functions, it is possible to construct a space with a finite number of points for which the t-function is the strongest.

On the other hand, the problem of finding the strongest possible t-function for a given SM-space is considerably more complicated, since it often depends on the particular form of the distribution function F_{pq}. Some general results along this line are developed at the end of this section. Also, it is shown that there are Menger spaces for which there is no strongest t-function.

Theorem 1. Let T be a t-function. Then there exists a Menger space for which T is the strongest t-function satisfying the postulate (SM-IV_m).

Proof: One can construct an SM-space such that for every point (a,b) in the unit square, there are points p,q,r in S and numbers x,y, such that $F_{pq}(x) = a$, $F_{qr}(y) = b$ and $F_{pr}(x + y) = T(F_{pq}(x),F_{qr}(x))$.

For a fixed (a,b) in the interior of the unit square, choose three points $\{p_1, p_2, p_3\}$. The distance distribution functions relating these three points are as follows:

$$F_{12}(x) = \begin{cases} 0, & x \leq 1 - \epsilon, \\ a, & 1 - \epsilon < x \leq 1, \\ 1, & 1 < x, \end{cases} \qquad F_{23}(x) = \begin{cases} 0, & x \leq 1 - \epsilon, \\ b, & 1 - \epsilon < x \leq 1, \\ 1, & 1 < x, \end{cases}$$

$$F_{13}(x) = \begin{cases} 0, & x \leq 2 - 2\epsilon, \\ T(a,b), & 2 - 2\epsilon < x \leq 2 - \epsilon, \\ 1, & 2 - \epsilon < x, \end{cases}$$

where $0 < \epsilon \leq 1/2$.

Using the results of Lemma 4 together with $0 < \epsilon \leq 1/2$, it follows that only the permutations of one triple of distinct points needs to be investigated. Furthermore, in view of the symmetry of F_{pq}, there are only three distinct permutations. Condition (SM-IV$_m$) is automatically satisfied for those values of x or y making either distribution function on the right zero (since $T(0,a) = 0$) and for those values of $x + y$ making the distribution function on the left unity (since $T(a,b) \leq 1$). The remaining possibilities need only to be checked.

Case I. To show $F_{13}(x + y) \geq T(F_{12}(x), F_{23}(y))$.

(a) Let $1 - \epsilon < x \leq 1$ and $1 - \epsilon < y \leq 1$. $F_{13}(x + y)$ is at least $T(a,b)$ and $T(F_{12}(x), F_{23}(y))$ is at most $T(a,b)$, then $F_{13}(x + y) \geq T(a,b) \geq T(F_{12}(x), F_{23}(y))$ implies $F_{13}(x + y) \geq T(F_{12}(x), F_{23}(y))$.

(b) For what follows it is important to note that equality is attained for $x = y = 1 - \epsilon/2$. Let $1 - \epsilon < x \leq 1$, $1 < y$, or $1 - \epsilon \leq y \leq 1$, $1 < x$. Then $2 - \epsilon < x + y$, so that $F_{13}(x + y) = 1$; hence, $F_{13}(x + y) \geq T(F_{12}(x), F_{23}(y))$.

Case II. To show $F_{12}(x + y) \geq T(F_{13}(x), F_{23}(y))$.

$F_{12}(x + y) = 0$ unless $x > 2 - 2\epsilon$ and $y > 1 - \epsilon$ in which case $F_{12}(x + y) = 1$. $F_{12}(x + y) = 1$, because $x + y \geq 3 - 3\epsilon$ or $x + y \geq 3(1 - \epsilon)$; therefore, $F_{12}(x + y) = 1$. On the other hand $F_{13}(x)$, where $x > 2 - 2\epsilon$, is at most one and $F_{23}(y)$ where $y > 1 - \epsilon$ is at most one; hence,

$$F_{12}(x + y) \geq T(F_{13}(x), F_{23}(y)).$$

Case III. To show $F_{23}(x + y) \geq T(F_{13}(x), F_{12}(y))$.

The left member is zero unless $x > 2 - 2\epsilon$ and $y > 1 - \epsilon$, hence the proof is similar to that of Case II.

Now with every point (a,b) in the interior of the unit square, a triple of points in S whose distance distribution functions are defined as above can be associated. Thus, each (a,b) has a triangle corresponding to it. Define the distribution function of any pair of points belonging to different triangles to be $H(x - 1)$, where H is the distribution function defined earlier. The totality of triangles of points so obtained will be the SM-space of the theorem.

It remains to show that (SM-IV$_m$) holds for distinct triples involving points from two or three triangles. If the three triangles are distinct, SM-IV$_m$ becomes $H(x + y - 1) \geq T(H(x - 1), H(y - 1))$. The right member is zero unless x and y both exceed one; but then $x + y > 1$, so that $H(x + y - 1) = 1$.

If there are two triangles, then one of them will contain exactly one point of the triple, say p. If p appears on the left in (SM-IV$_m$), then

$x + y > 1$ makes the left member one and $x + y \leq 1$ makes the right member zero. (Note: This is why one needs to take $\epsilon \leq 1/2$.) If p does not appear in the left member of (SM-IV_m), then it appears in both distribution functions on the right. Hence the right member is zero unless x and y are greater than 1, but then $x + y > 2$ and the left member is one.

It has thus been established that the collection of triangles is a Menger space under the given T. Furthermore, for this Menger space, the given T is the strongest possible. It cannot be strengthened at any point (a,b) in the interior of the square; for, by construction, the corresponding triangle is such that for some numbers x,y (SM-IV_m) holds with equality when the argument of T is (a,b).

Also, it should be noted that since the associativity of t-functions was not used in the proof of the theorem, the results will be true for functions which are not strict t-functions.

In constraining T so that it cannot be strengthened, the construction in the proof of Theorem 1 used distinct triangles for distinct points in the domain of T. There is, therefore, an uncountable number of points in the resulting Menger space. However, a space with a countable number of points for which T is continuous will be sufficient; since it suffices to constrain T on a countably dense subset of the unit square. [7]

The question of whether a given Menger space has a strongest T such that (SM-IV_m) holds is the next item to be investigated.

Theorem 2. If $(S,\bar{F})$ is a given SM-space and $\{T_\alpha\}$, α in A, is a collection of functions, each satisfying (a'), (b), (c), and (SM-IV_m) on the space $(S,\bar{F})$, then so does $\sup T_\alpha$.

Proof: Condition $T(0,0) = 0$, $T(a,1) = T(1,a) = a$ implies that $\{T_\alpha\}$ is bounded. The collection $\{T_\alpha\}$ being bounded implies that sup T_α satisfies (a'), (b), (c), and ($SM\text{-}IV_m$).

Corollary. For every Menger space there is a unique strongest T satisfying all the conditions required of a t-function except possibly the associativity condition.

Example: Let $T_1(x,y) = \text{Max}(x + y - 1,0)$. Let $T_2(x,y) = 3/4$ if $3/4 \leq x,y < 1$, $T_2(a,1) = T_2(1,a) = a$, and $T_2(x,y) = 0$ otherwise. Both T_1 and T_2 can be shown to be associative. However, $T = \text{Max}(T_1,T_2)$ is not associative because $T(T(3/4,3/4),1/2) = T(3/4,1/2) = 1/4$ while $T(3/4,T(3/4,1/2))= T(3/4,1/4) = 0$.

This example shows that the supremum of two associative functions need not be associative and leads at once to the next theorem.

Theorem 3. There is a Menger space for which there is no strongest t-function.

Proof: Using Theorem 2, the T in example 2 has all the properties of a t-function except associativity. It was noted immediately following the proof of Theorem 1, that the construction in the theorem applies to yield an SM-space such that this T is the strongest function satisfying (a'), (b), (c), and ($SM\text{-}IV_m$). This SM-space is a Menger space under T_1 and T_2. Thus, if there were a strongest t-function, it would be stronger than T. Since this is impossible, the proof is complete.

At present it is not known whether the set of allowable t-functions for a Menger space will have a maximal element. A partial result can be obtained in this direction.

Definition 4. A two-place function F is left continuous if $\lim_{x \to a-} F(x,b) = F(a,b)$ and $\lim_{y \to b-} F(a,y) = F(a,b)$ for all (a,b) in the domain of F.

Theorem 4. If $(S,\overline{F})$ is an SM-space and G the set of all left-continuous t-functions T for which $(S,\overline{F},T)$ is a Menger space, then G has a maximal element.

Proof: The set G is partially ordered by the relation "stronger than." Since a partially ordered set contains at least one maximal simply ordered subset (Zorn's Lemma), it is sufficient to show that the supremum of any totally ordered subset of G is again in G. Let $\{T_\alpha\}$, α in A, be such a totally ordered subset. The order on $\{T_\alpha\}$ induces an order $(\leq)$ on A, where $\alpha \leq \beta$ if and only if $T_\alpha \leq T_\beta$.

Let $T = \sup T_\alpha$. In view of Theorem 2 it suffices to show that T is left continuous and associative.

Part I. To show that T is left continuous, let x,y, $0 \leq x,y \leq 1$, and $\epsilon > 0$ be given. Then there is an α such that $0 \leq T(x,y) - T_\alpha(x,y) < \epsilon/2$; and because of the left-continuity and monotonicity of T_α, there is a $\delta > 0$ such that for $0 < x - x' < \delta$,

$$0 \leq T_\alpha(x,y) - T_\alpha(x',y) < \epsilon/2.$$

Furthermore, from the definition of T,

$$T_\alpha(x',y) - T(x',y) < 0.$$

Combining these three inequalities yields,

$$T(x,y) - T(x',y) < \epsilon,$$

whenever $0 < x - x' < \delta$; hence, in view of the monotonicity and symmetry, T is left-continuous.

Part II. To show that T is associative, it suffices to show that $T(T(a_1,b_1)c_1)$ is invariant under any permutation of the elements (a_1,b_1,c_1). Let (a_2,b_2,c_2) be such a permutation and let $\epsilon > 0$ be given. Since $\{T_\alpha\}$ is totally ordered, there is a β such that

$$0 \leq T(T(a_i,b_i),c_i) - T_\beta(T(a_i,b_i),c_i) < \epsilon/2, \qquad i = 1,2$$

and, from the definition of T and the fact that T_β is left-continuous, there is an $\alpha \geq \beta$ such that both $T(a_1,b_1) - T_\alpha(a_1,b_1)$ and $T(a_2,b_2) - T_\alpha(a_2,b_2)$ are so small that

$$0 \leq T_\beta(T(a_i,b_i),c_i) - T_\beta(T_\alpha(a_i,b_i),c_i) < \epsilon/2, \qquad i = 1,2.$$

Since $\alpha \geq \beta$, it follows that

$$T_\beta(T_\alpha(a_i,b_i),c_i) - T_\alpha(T_\alpha(a_i,b_i),c_i) \leq 0, \qquad i = 1,2.$$

Combining the inequalities one obtains

$$0 \leq T(T(a_1,b_1),c_1) - T_\alpha(T_\alpha(a_1,b_1),c_1) < \epsilon,$$

and

$$0 \leq T(T(a_2,b_2),c_2) - T_\alpha(T_\alpha(a_2,b_2),c_2) < \epsilon.$$

Now T_α is associative and symmetric, so that

$$T_\alpha(T_\alpha(a_1,b_1),c_1) = T_\alpha(T_\alpha(a_2,b_2),c_2).$$

Consequently,

$$T(T(a_1,b_1),c_1) - T(T(a_2,b_2),c_2) < \epsilon;$$

hence the proof is now complete.

3. Wald Space. The other generalized triangle inequality to be considered is due to A. Wald. It is

$$\mathrm{IV}_W \quad F_{pr}(x) \geq \left[F_{pq} * F_{qr}\right](x), \quad \text{for all } x \geq 0,$$

where $*$ is the convolution defined by

$$\left[F_{pq} * F_{qr}\right](x) = \int_{-\infty}^{\infty} F_{pq}(x - y)\, dF_{qr}(y).$$

Since $F_{pq}(x - y) = 0$ for $y \geq x$ and $F_{qr}(y) = 0$ for $y \leq 0$, it follows that

$$\left[F_{pq} * F_{qr}\right](x) = \int_{0}^{x} F_{pq}(x - y)\, dF_{qr}(y).$$

Due to the fact that the convolution of the distribution functions of two independent random variables gives the distribution of their sum, the interpretation of IV_W is: The probability that the distance from p to r is less than x is greater than or equal to the probabilities that the distance from p to q and the distance from q to r is less than x.

Definition 5. A Wald space is an SM-space in which IV_W holds universally.

Theorem 5. A Wald space is a Menger space under the choice T = Product.

Proof: Choosing a Wald space and for any $x,y \geq 0$

$$F_{pr}(x+y) \geq \int_0^{x+y} F_{pq}(x+y-z)\,dF_{qr}(z),$$

$$\int_0^{x+y} F_{pq}(x+y-z)\,dF_{qr}(z) = \int_0^{x+y}\left[\int_0^{x+y-z} dF_{pq}(t)\right] d\,F_{qr}(z)$$

$$= \iint\limits_{\substack{t,z \geq 0 \\ t+z \leq x+y}} dF_{pq}(t)\,dF_{qr}(z),$$

$$\iint\limits_{\substack{t,z > 0 \\ t+z \leq x+y}} dF_{pq}(t)\,dF_{qr}(z) \geq \iint\limits_{\substack{0 \leq t \leq x \\ 0 \leq z \leq y}} dF_{pq}(t)\,dF_{qr}(z).$$

Since the rectangle $\{(t,z);\ 0 \leq t \leq x,\ 0 \leq z \leq y\}$ is contained in the rectangle $\{(t,z)\ t,z \geq 0\}$ and the F's are non-decreasing

$$\iint\limits_{\substack{0\leq t\leq x\\ 0\leq z\leq y}} dF_{pq}(x)dF_{qr}(z) = \int_0^x\int_0^y dF_{pq}(t)dF_{qr}(z)$$

$$= \int_0^x dF_{pq}(t)\int_0^y dF_{qr}(z) = F_{pq}(x)F_{qr}(y).$$

Hence $F_{pr}(x + y) \geq F_{pq}(x)F_{qr}(x)$.

Corollary. If the Wald inequality holds, then so does the inequality IV.

Proof: If $F_{pq}(x) = 1$ and $F_{qr}(y) = 1$, then by using the above theorem $F_{pr}(x + y) \geq 1$; but $F_{pr}(x + y)$ cannot be greater than one by definition, hence $F_{pr}(x + y) = 1$.

Lemma 5. If the points p,q,r are not all distinct, then IV_W holds for the triple p,q,r.

Case I. If $p = r \neq q$,

then $F_{pr}(x) = 1$ for $x > 0$ and $F_{pr}(x) = 0$ for $x > 0$; hence $F_{pr}(x) = H$ and, therefore, $F_{pr}(x) \geq \int_0^x F_{pq}(x - y)dF_{qr}(y)$.

Case II. If $p = q \neq r$, then for $x \geq 0$

$$F_{pr}(x) = F_{qr} = \int_0^x dF_{qr}(y) = \int_0^x H(x - y)dF_{qr}(y)$$

$$= \int_0^x F_{pp}(x - y)dF_{qr}(y) = \int_0^x F_{pq}(x - y)dF_{qr}(y).$$

Case III. If $p \neq q = r$, then for $x \geq 0$

$$F_{pr}(x) = F_{pq}(x) = \int_0^x dF_{pq}(y) = \int_0^x H(x-y)d(F_{pq}(y))$$

$$= \int_0^x F_{qq}(x - y)d(F_{pq}(y)) = \int_0^x F_{qr}(x - y)dF_{pq}(y).$$

Theorem 6. If in an SM-space IV_m holds under T = Max for all triples of distinct points, then the space is a Wald space.

Proof: Let p,q,r be distinct points. Then for any $x \geq 0$

$$F_{pq}(x) \geq \mathrm{Max}(F_{pq}(0), F_{qr}(x)) = F_{qr}(x) = \int_0^x dF_{pr}(y)$$

$$\geq \int_0^x F_{pq}(x - y)dF_{qr}(x) \text{ since } 0 \leq F_{pq}(x - y) \leq 1.$$

Hence IV_W holds for all triples of distinct points in the space. Using the preceding lemma, IV_W holds automatically for all triples of non-distinct points. Consequently, IV_W holds for all triples of points in the space.

4. Equilateral spaces. The simplest metric spaces are called equilateral spaces. Where the distance between p and q is

$$d(p,q) = \begin{cases} a & \text{if } p \neq q \\ 0 & \text{if } p = q \end{cases}$$

Now an SM-space is equilateral if for some distribution function G satisfying $G(0) = 0$,

$$F_{pq}(x) = \begin{cases} G(x) & \text{if } p \neq q \\ H(x) & \text{if } p = q \end{cases}$$

where H is the distribution function described earlier.

Theorem 7. The means, medians, etc. of the statistical distances in an equilateral SM-space form an M-space.

Proof: Any one of the quantities is zero when p = q and a fixed positive number for any p,q when $p \neq q$. Hence they satisfy the postulates for a metric space.

Theorem 8. In an equilateral SM-space, the Menger triangle inequality IV_m holds for any triple of distinct points under T = Max and holds universally under T = Min.

Proof: $G(x + y) \geq \text{Max}(G(x),G(y)) \geq \text{Min}(G(x),G(y))$ holds when all three points are distinct. When all three are not distinct, the result is

$$G(x + y) \geq \text{Min}(G(x),1).$$

Corollary. An equilateral SM-space is a Wald space.

Proof: By the above theorem it should be noted that an equilateral SM-space is a Menger space where T = Min holds universally. Hence using Theorem 2, it follows that an equilateral SM-space is also a Wald space.

Examples of equilateral SM-spaces in which IV_m holds for stronger choices of T follows.

Example 1.

$$G(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 \leq x \leq 1 \\ 1 & 1 \leq x \end{cases}$$

This is a very simple form of the uniform distribution function. It is noted that for any triple of distinct points in this space IV_m holds under T = Min(Sum,1), since $G(x + y) \geq \text{Min}(G(x) + G(y),1)$.

Example 2.

$$G(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-x} & x \geq 0 \end{cases}$$

For any triple of distinct points in this space, IV_m holds under T = Sum - Product. To prove the previous statement, it suffices to consider values of $x,y \geq 0$. Thus, the t-function becomes

$$T(G(x),G(y)) = 1 - e^{-x} + 1 - e^{-y} - 1 + e^{-x} + e^{-y} - e^{-(x + y)}$$
$$= 1 - e^{-(x + y)}.$$

Using the above results along with the definition of $G(x)$, it follows that

$$G(x + y) \geq \text{Sum} - \text{Product}.$$

The above two examples show that there are equilateral SM-spaces in which the generalized triangle inequality IV_m holds under a stronger T than T = Max. In order to show that T = Max is the best possible t-function, the following example will suffice.

Example 3.

$$G(x) = \begin{cases} 0, & x \leq 0 \\ a, & 0 < x \leq k \\ b, & k < x \leq 3k \\ 1, & 3k < x \end{cases}$$

where $a \leq 1$, $0 < a < b \leq 1$, and k is any positive number. Then for $0 < x \leq k$, $k < y \leq 3k$, $G(x + y) = b = \text{Max}(a,b)$; hence IV_m cannot hold under a choice of T which is stronger than T = Max.

5. Simple Spaces. Let (S,d) be an M-space and G a distribution function, differing from H, and satisfying $G(0) = 0$. For each pair of points p,q in S, define the distribution function F_{pq} as follows:

$$F_{pq}(x) = \begin{cases} G(x/d(p,q)), & p \neq q \\ H(x) & p = q. \end{cases}$$

Definition 6. An SM-space $(S,\overline{F})$ is said to be a simple space if and only if there exist a metric d on S and a distribution function G satisfying $G(0) = 0$; such that, for every pair of points p,q in S, $\overline{F}(p,q) = F_{pq}$ is given by

$$F_{pq}(x) = \begin{cases} G(x/d(p,q)), & p \neq q \\ H(x) & p = q \end{cases}$$

Furthermore, $(S,\overline{F})$ is called the simple space generated by the M-space (S,d) and the distribution function G.

Theorem 9. A simple space is a Menger space under any choice of T satisfying (a'), (b), (c), and (d').

Proof: It is sufficient to show that IV_m holds under T = Min since this is the strongest T possible. Thus, in view of Lemma 4, it suffices to show that for p,q,r distinct

$$(1) \qquad G\frac{(x+y)}{d(p,r)} \geq \mathrm{Min}(G(x/d(p,q)),G(y/d(p,r))).$$

Note: $d(p,r) \leq d(p,q) + d(q,r)$ implies that $\dfrac{1}{d(p,r)} \geq \dfrac{1}{d(p,q) + d(q,r)}$,

hence
$$(2) \qquad \frac{(x+y)}{d(p,r)} \geq \frac{x+y}{d(p,q)+d(q,r)}.$$

Also, since d(p,r) and d(p,q) are positive,

$$(3) \qquad \mathrm{Max}(x/d(p,q),y/d(q,r)) \geq \frac{x+y}{d(p,q)+d(q,r)}$$

$$\geq \mathrm{Min}(x/d(p,q),y/d(q,r))$$

with equality following if and only if $x/d(p,q) = y/d(q,r)$. On combining the inequality (2) and the right-hand inequality in (3), one has

$$\frac{x+y}{d(p,r)} \geq \mathrm{Min}(x/d(p,q),y/d(q,r)),$$

which, since G is non-decreasing, implies

$$G\left(\frac{x+y}{d(p,r)}\right) \geq \mathrm{Min}\ G(x/d(p,q)),G(y/d(q,r));$$

hence the theorem follows.

Corollary. An equilateral M-space generates an equilateral SM-space.

Proof: Now

$$d(p,q) = \begin{cases} a & \text{if } p \neq q \\ 0 & \text{if } p \neq q. \end{cases}$$

Then

$$F_{pq}(x) = \begin{cases} G(x/a) = G(X) & p \neq q \\ H(x) \quad = H(x) & p = q. \end{cases}$$

From the left-hand member of (3) where T = Max for distinct triples of points p,q,r such that $d(p,r) \leq d(p,q) + d(q,r)$, it follows that IV_m may not hold universally. This leads to the following theorem.

Theorem 10. If (S,d) is a finite-dimensional Euclidean space, G is a continuous distribution function such that $G(0) = 0$ and $0 < G(x) < 1$ for all $x > 0$; then Min is the strongest T under which IV_m holds for all triples of distinct points.

Proof: Suppose T is stronger than Min. Then there exists at least one pair of numbers $a,b (0 < a,b < 1)$, such that $T(a,b) > \mathrm{Min}(a,b)$.

Case I. If $a = b$, choose $x = y$ such that $G(x) = a$ and choose $d(p,q) = d(q,r) = 1, d(p,r) = 2$. Then since equality is attained in inequality (1), one cannot have $T(a,a) > \mathrm{Min}(a,a) = a$.

Case II. If $a \neq b$, one can choose $a < b$. Let $\epsilon = T(a,b) - \mathrm{Min}(a,b)$ and let u,v, be such that $a = G(u)$ and $b = G(v)$. Such numbers u and v clearly exist since G is a continuous distribution function; in addition, $u < v$.

Moreover, since G is continuous, there exists an $h > 0$ such that

$$G(u + h) < G(u) + \epsilon = a + \epsilon .$$

Let $d(q,r) = t$ be fixed. Then $d(p,q)$ can be chosen so that

$$d(p,q) = s \text{ such that } \frac{t}{s + t} < \frac{h}{v - u},$$

$$d(p,r) \leq d(p,q) + d(q,r) = s + t,$$

$$x = ud(p,q) \text{ and } y = vd(q,r).$$

Then

$$\mathrm{Min}\left[G(x/d(p,q)), G(y/d(q,r))\right] = \mathrm{Min}\left[G(u), G(v)\right] = \mathrm{Min}(a,b) = a;$$

furthermore,

$$G\left(\frac{x + y}{d(p,r)}\right) = G\left(\frac{us + vt}{s + t}\right) = G\left(u + \frac{(v - u)}{s + t}\, t\right) \leq G(u + h) < a + \epsilon .$$

This contracts the hypothesis

$$G\left(\frac{x+y}{d(p,r)}\right) = T(a,b) \geq \mathrm{Min}(a,b) + \epsilon = a + \epsilon .$$

Up to this point it has been shown that every Wald space is a Menger space under T = Product, and also that every equilateral space is a Wald space. This is not the case for simple spaces as the results of the next theorem will indicate.

Theorem 11. There exist simple spaces which are not Wald spaces.

Proof: A counter-example will be sufficient.

Example. Let the distribution function be defined by:

$$(4) \quad F_{pq}(x) = 1 - e^{-x/d(p,q)}.$$

With $d(p,q) = R, d(q,r) = S$ and $d(p,r) = T$, two cases arise.

Case I.

$$\left[F_{pq} * F_{qr}\right](x) = \int_0^x F_{pq}(x-y)\,dF_{qr}(y) \qquad R \neq S$$

$$= \int_0^x 1 - \exp - \frac{(x-y)}{R}\, d(1 - \exp(-\tfrac{y}{S}))$$

$$= \int_0^x (1 - \exp - \frac{(x-y)}{R}\, \frac{1}{S} \exp(-\tfrac{y}{S}))dy$$

$$= \frac{1}{S}\int_0^x (\exp(-\tfrac{y}{S}) - \exp(-\tfrac{x}{R})\exp(-\tfrac{y}{S} + \tfrac{y}{R}))dy$$

$$= \frac{1}{S}\left(\int_0^x \exp(-\tfrac{y}{S})dy - \exp(-\tfrac{x}{R})\int_0^x \exp(-\tfrac{R+S}{SR}y)dy\right)$$

$$= 1 - \frac{1}{R-S}\, R \cdot \exp(-\tfrac{x}{R}) - S \cdot \exp(-\tfrac{x}{S}).$$

Case II.

$$\left[F_{pq} * F_{qr}\right](x) = \int_0^x F_{pq}(x - y)\,dF_{qr}(y) \qquad R = S$$

$$= \int_0^x 1 - \exp(-\frac{(x - y)}{R})\,d(1 - \exp(-\frac{y}{R}))$$

$$= \frac{1}{R}\int_0^x (\exp(-\frac{y}{R}) - \exp(-\frac{x}{R}))\,dy$$

$$= \left[1 - \exp(-\frac{x}{R}) - \frac{1}{R}\cdot\exp(-\frac{x}{R})x\right]$$

$$= 1 - (1 + \frac{x}{R})\exp(-\frac{x}{R}).$$

The above results can be stated briefly as:

$$(5)\qquad \left[F_{pq} * F_{qr}\right](x) = \begin{cases} 1 - \frac{1}{R - S}(R \cdot \exp(-\frac{x}{S})) & R \neq S \\ 1 - (1 + \frac{x}{R}) \cdot \exp(-\frac{x}{R}) & R = S. \end{cases}$$

In order that the Wald inequality IV_W be satisfied, it must be true that

$$(6)\qquad F_{pq}(x) \geq \left[F_{pq} * F_{qr}\right](x) \text{ for every } x > 0.$$

Suppose $R \neq S$, say $R > S$. Then keeping x fixed and applying the mean value theorem to the second term of

$$\left[F_{pq} * F_{qr}\right](x) = 1 - \frac{1}{R - S}(R \cdot \exp(-\frac{x}{R}) - S \cdot \exp(-\frac{x}{S})),$$

for $R \neq S$ it follows that

$$(7)\qquad \left[F_{pq} * F_{qr}\right](x) = 1 - (1 + \frac{x}{t})\exp(-\frac{x}{t})), \text{ where } S < t < R.$$

Furthermore, if $R = S$, it follows on comparing (7) with (4) that (7) holds when $t = R$. Thus, in both cases in order that (6) holds, it is necessary that

$$(1 + \frac{x}{t})\exp(-\frac{x}{t}) \geq \exp(-\frac{x}{T}) \qquad \text{for all } x > 0;$$

that is

$$1 + \frac{x}{t} \geq \exp(x(\frac{1}{t} - \frac{1}{T}) \qquad \text{for all } x > 0.$$

This is true if and only if $(\frac{1}{t} - \frac{1}{T}) \leq 0$; therefore, $T \leq t$. In particular it is necessary that $T \leq R$. This means that the side of the triangle pqr whose length is T certainly cannot be the longest side of the triangle. Thus, it must be concluded: If $d(p,r) \geq \text{Max}(d(p,q,d(q,r))$, then the Wald inequality will fail to hold for sufficiently large x.

TOPOLOGY, CONVERGENCE, CONTINUITY

In order to discuss the topology of an M-space, the definition of a neighborhood is needed.

Definition. A set U in a topological space $(X,\mathcal{J})$ is a neighborhood ($\mathcal{J}$--neighborhood) of a point x if and only if U contains an open set to which x belongs.

A similar definition for a neighborhood in an SM-space is needed before one can examine the topology of an SM-space. There are several non-equivalent definitions for a neighborhood in an SM-space. The one defined here most closely resembles the classical neighborhood on an M-space.

Definition 7. Let p be a point in the SM-space $(S,\overline{F})$. By an ϵ, λ-neighborhood of p, $\epsilon > 0$, $\lambda > 0$ is meant the set of all points q in S in which $F_{pq}(\epsilon) > 1 - \lambda$. Thus,

$$N_p(\epsilon, \lambda) = \left\{q; F_{pq}(\epsilon) > 1 - \lambda\right\}.$$

The interpretation is: $N_p(\epsilon, \lambda)$ is the set of all points q in S for which the probability of the distance from p to q being less than ϵ is greater than $1 - \lambda$. Observe that the above definition shows that the neighborhood of a point in an SM-space depends on two parameters.

Theorem 12. In a simple space, $N_p(\epsilon, \lambda)$ is an ordinary spherical neighborhood of p in the generating M-space.

Proof: For any p,q it follows that

$$F_{pq}(\epsilon) = G(\epsilon/d(p,q)),$$

which will be greater than $1 - \lambda$ provided only that d(p,q) is sufficiently small.

Lemma 5. If $\epsilon_1 \leq \epsilon_2$ and $\lambda_1 \leq \lambda_2$, then $N_p(\epsilon_1, \lambda_1) \subset N_p(\epsilon_2, \lambda_2)$.

Proof: Suppose $q \in N_p(\epsilon_1, \lambda_1)$, so that $F_{pq}(\epsilon_1) > 1 - \lambda_1$. Then $F_{pq}(\epsilon_2) \geq F_{pq}(\epsilon_1) > 1 - \lambda_1 > 1 - \lambda_2$; hence by definition $q \in N_p(\epsilon_2, \lambda_2)$.

Theorem 13. If (S,$\mathcal{F}$) is a Menger space and T is continuous, then (S,$\mathcal{F}$) is a Hausdorff space in the topology induced by the family of ϵ, λ- neighborhoods $\{N_p\}$.

Proof: In order to prove the above theorem, the following four conditions must be satisfied.

(A) For every p in S, there exists at least one neighborhood, N_p of p and every neighborhood of p contains p.

(B) If N_p^1 and N_p^2 are neighborhoods of p, then there exists a neighborhood of p, N_p^3 such that $N_p^3 \subset N_p^1 \cdot N_p^2$. (Note: $N_p^1 \cdot N_p^2$ is used to represent the intersection of the two neighborhoods.)

(C) If N_p is a neighborhood of p and $q \in N_p$, then there exists a neighborhood of q, N_q, such that $N_q \subset N_p$.

(D) If $p \neq q$, then there exist disjoint neighborhoods N_p and N_q; such that p is an element of N_p and q is an element of N_q.

Proof:

(A) For every $\epsilon > 0$ and every $\lambda > 0$, p is an element of $N_p(\epsilon, \lambda)$ since $F_{pp}(\epsilon) = 1$ for any $\epsilon > 0$.

(B) Let

$$N_p^1(\epsilon_1, \lambda_1) = \{q; F_{pq}(\epsilon_1) > 1 - \lambda_1\}$$

and

$$N_p^2(\epsilon_2, \lambda_2) = \{q; F_{pq}(\epsilon_2) > 1 - \lambda_2\}$$

be the given neighborhoods of p, and consider

$$N_p^3 = \{q; F_{pq}(\mathrm{Min}(\epsilon_1, \epsilon_2)) > 1 - \mathrm{Min}(\lambda_1, \lambda_2)\} .$$

Clearly p is an element of N_p^3; and since $\mathrm{Min}(\epsilon_1, \epsilon_2) \leq \epsilon_1$ and $\mathrm{Min}(\lambda_1, \lambda_2) \leq \lambda_1$, by Lemma 5, $N_p^3 \subset N_p^1$. Similarly $N_p^3 \subset N_p^2$, hence $N_p^3 \subset N_p^1 \cdot N_p^2$.

(C) Let $N_p = \{r; F_{pr}(\epsilon_1) > 1 - \lambda_1\}$ be the given neighborhood of p. Since $q \in N_p$

$$F_{pq}(\epsilon_1) > 1 - \lambda_1.$$

Now F_{pq} is left continuous at ϵ_1. Hence, there exist an $\epsilon_0 < \epsilon_1$ and a $\lambda_0 < \lambda_1$ such that

$$F_{pq}(\epsilon_0) > 1 - \lambda_0 > 1 - \lambda_1.$$

Let $N_q = \{r; F_{pr}(\epsilon_2) > 1 - \lambda_2$, where $0 < \epsilon_2 < \epsilon_1 - \epsilon_0$ and λ_2 is chosen such that

$$T(1 - \lambda_0, 1 - \lambda_2) > 1 - \lambda_1.$$

Such a λ_2 exists since by hypothesis T is continuous, $T(a,1) = a$, and $1 - \lambda_0 > 1 - \lambda_1$. Now suppose $s \in N_q$, so that

$$F_{qs}(\epsilon_2) > 1 - \lambda_2.$$

Then

$$F_{ps}(\epsilon_1) \geq T(F_{pq}(\epsilon_0), F_{qs}(\epsilon_1 - \epsilon_0)) \geq T(F_{pq}(\epsilon_0), F_{qs}(\epsilon_2)) \geq T(1 - \lambda_0, 1 - \lambda_2) > 1 - \lambda_1.$$

This means s is an element of N_p; hence, N_q is contained in N_p.

(D) Let $p \neq q$. Then there exist real numbers x and a, $x > 0$, $0 \leq a < 1$, such that $F_{pq}(x) = a$. Let

$$N_p = \{r; F_{pr}(x/2) > b\} \text{ and } N_q = \{r; F_{qr}(x/2) > b\},$$

where b is chosen so that $0 < b < 1$ and $T(b,b) > a$. Such a number b exists, since T is continuous and $T(1,1) = 1$. Now suppose there is a point s in $N_p \cdot N_q$ so that $F_{ps}(x/2) > b$ and $F_{qs}(x/2) > b$. Then

$$a = F_{pq}(x) \geq T(F_{ps}(x/2), F_{qs}(x/2)) \geq T(b,b) > a$$

which is a contradiction. Thus, N_p and N_q are disjoint.

The concept of being a Hausdorff space in the topology induced by the family of ϵ, λ-neighborhood $\{N_p\}$ will be investigated further; but first, continuity properties in an SM-space will be considered.

In an M-space the notion of convergence of a sequence of points $\{p_n\}$ to a point p is introduced by using the concept of neighborhoods. A distance function d is continuous on S if $p_n \rightarrow p$ and $q_n \rightarrow q$, implies $d(p_n,q_n) \rightarrow d(p,q)$.

Using the definition of neighborhood in an SM-space, one finds that a significant difference arises, for there are two distinct types of questions regarding convergence and continuity to be investigated, namely,

(a) Those relating to the distance function $\overline{F}(x)$ considered d as a function on $S \times S$--either for a fixed value of x or for a range of values.

(b) Those relating to the individual distribution functions F_{pq}--either for a fixed pair of points (p,q) or for a set of pairs of points.

As was to be expected, these questions are not independent.

<u>Definition 8</u>. A sequence of points $\{p_n\}$ in an SM-space is said to <u>converge</u> to a point p in S (denoted by $p_n \longrightarrow p$) if and only if for every $\lambda > 0$ and every $\epsilon > 0$, there exists an integer $M_{\epsilon,\lambda}$, such that p_n is an element of $N_p(\epsilon, \lambda)$; therefore, $F_{pq_n}(x) = 1 - \lambda$ whenever $n > M_{\epsilon,\lambda}$.

Lemma 6. If $p_n \longrightarrow p$, then $F_{pp_n} \longrightarrow F_{pp} = H$.

Proof: (a) if $x > 0$, then for every $\lambda > 0$ there exists an integer $M_{x,\lambda}$ such that $F_{pp_n}(x) > 1 - \lambda$ whenever $n > M_{x,\lambda}$. This means that $\lim_{n \to \infty} F_{pp_n}(x) = 1 = F_{pp}(x)$.

(b) If $x = 0$, then for every n, $F_{pp_n}(0) = 0$; and hence, $\lim_{n \to \infty} F_{pp_n}(0) = 0 = F_{pp}(0)$.

If $F_{pp_n} \longrightarrow F_{pp}$, then for every $\epsilon > 0$ and every $\lambda > 0$ there exists an integer $M_{\epsilon,\lambda}$ such that p_n is an element of $N_p(\epsilon, \lambda)$; hence by definition $p_n \longrightarrow p$.

Corollary. The convergence is uniform on any closed interval $[a,b]$ such that $a > 0$; therefore, $M_{x,\lambda}$ is independent of x for $a \leq x \leq b$, $a > 0$.

Proof: For any x, $a \leq x \leq b$, $(a > 0)$ $F_{pp_n}(x) \geq F_{pp_n}(a)$.

Theorem 14. If $(S, \overline{F})$ is a Menger space and T is continuous, then the statistical distance function, $\overline{F}$, is a lower semi-continuous function of points; therefore, for every fixed x, if $q_n \longrightarrow p$ and $p_n \longrightarrow p$, then $\liminf_{n \to \infty} F_{p_n q_n}(x) = F_{pq}(x)$.

Proof: If $x = 0$, then for every n, $F_{p_n q_n}(0) = 0 = F_{pq}(0)$. Suppose then that $x > 0$, and let $\epsilon > 0$ be given. Since F_{pq} is left-continuous at x, there is an h, $0 < 2h < x$, such that

$$F_{pq}(x) - F_{pq}(x - 2h) < \epsilon/3.$$

Set $F_{pq}(x - 2h) = a$. Since T is continuous and $T(a,1) = a$, there is a number t, $0 < t < 1$, such that

$$T(a,t) > a - \epsilon/3$$

and

$$T(a - \epsilon/3, t) > a - 2\epsilon/3.$$

Since $q_n \longrightarrow q$ and $p_n \longrightarrow p$ by Lemma 6, there exists an integer $M_{h,t}$ such that $F_{qq_n}(h) > t$ and $F_{pp_n}(h) > t$, whenever $n > M_{h,t}$.

Now

$$F_{p_nq_n}(x) \geq T(F_{pq}(x - h), F_{pp_n}(h)),$$

and

$$F_{p_nq}(x - h) \geq T(F_{pq}(x - 2h), F_{pp_n}(h)).$$

Thus on combining the various inequalities, it follows that

$$F_{p_nq}(x - h) \geq T(a,t) > a - \epsilon/3.$$

Hence,

$$F_{p_nq_n}(x) \geq T(a - \epsilon/3, t) > a - 2\epsilon/3 > F_{pq}(x) - \epsilon .$$

Corollary 1. Let p be a fixed point and suppose $q_n \longrightarrow q$. Then

$$\liminf_{n \longrightarrow \infty} F_{pq_n}(x) = F_{pq}(x).$$

Corollary 2. If $(S,\overline{F})$ is a Wald space, then $\overline{F}$ is a lower semicontinuous function of points.

Proof: By Theorem 5, the Menger inequality holds in a Wald space under the continuous function T = Product.

Theorem 15. Let $(S,\overline{F})$ be a Menger space. Let T be continuous and at least as strong as Max(Sum - 1,0). Let $p_n \longrightarrow p$, $q_n \longrightarrow q$ and assume that F_{pq} is continuous at x. Then $F_{p_nq_n}(x) \longrightarrow F_{pq}(x)$. Therefore, the distance function $\overline{F}$ is a continuous function of points at (p,q,x), or expressed in another way, the sequence of functions $\{F_{p_nq_n}\}$ converges weakly to F_{pq}.

Proof: Using Theorem 14, it is sufficient to prove that $F_{p_nq_n}$ is upper semicontinuous; that is for $\epsilon > 0$ and n sufficiently large,

$$F_{p_nq_n}(x) < F_{pq}(x) + \epsilon.$$

Suppose then that $\epsilon > 0$ is given. Since F_{pq} is continuous at x, there exists an $h > 0$, such that

$$F_{pq}(x + 2h) - F_{pq}(x) < \epsilon/3.$$

By Lemma 6, there is an integer M such that the conditions,

$$F_{pp_n}(h) > 1 - \epsilon/3,$$

$$F_{qq_n}(h) > 1 - \epsilon/3,$$

are simultaneously satisfied for all $n > M$. From IV_m,

$$F_{pq}(x + 2h) \geq T(F_{pq_n}(x + h), F_{qq_n}(h))$$

and

$$F_{pq_n}(x + h) > T(F_{pq_n}(x), F_{pp_n}(h)).$$

Now by the hypotheses, T is at least as strong as Max(Sum - 1,0), so that by combining the inequalities, it follows that

$$F_{pq_n}(x + h) \geq F_{p_nq_n}(x) + F_{pp_n}(h) - 1 > F_{p_nq_n}(x) - \epsilon/3;$$

also

$$F_{pq}(x + 2h) \geq F_{pq_n}(x + h) + F_{qq_n}(h) - 1 > F_{p_nq_n}(x) - 2\epsilon/3.$$

Upon combining this last inequality with

$$F_{pq}(x + 2h) - F_{pq}(x) < \epsilon/3,$$

$$F_{p_nq_n}(x) < F_{pq}(x) + \epsilon.$$

This completes the proof.

Corollary 1. Under the hypotheses of Theorem 15, if $q_n \longrightarrow q$, then $F_{pq_n}(x) \longrightarrow F_{pq}(x)$.

Corollary 2. If the functions F_{pq} are each continuous functions for all p, q, in S, then $\overline{F}$ is a continuous function of points.

Corollary 3. If $(S,\overline{F})$ is a Wald space and if the functions F_{pq} are each continuous, then $\overline{F}$ is a continuous function of points. There are several examples which show that, given a topological space S, it is not true that one can always define a metric for S which will induce the given topology of S. Making S into a metric set is a simple matter; for example, a metric can be defined for S by

$$d(p,q) = 1 \qquad \text{if } p \neq q,$$

$$d(p,q) = 0 \qquad \text{if } p = q.$$

Obtaining a metric for S that will induce the original topology of S is a far deeper problem. Such a metric may be obtained only if the given space S is what is known as a metrizable space.

Definition 9. Let S be a topological space with topology $\mathcal{T}$. Then S is said to be metrizable if and only if it is possible to define a metric for S which will induce the topology $\mathcal{T}$. In particular, a metric space is a metrizable space for which a metric has been specified.

It has been shown that a large number of SM-spaces are Hausdorff spaces. As an immediate consequence of a generalization of Theorem 13, one can prove that a large number of SM-spaces are metrizable. Therefore, in numerous instances the existence of a statistical metric implies the existence of an ordinary metric.

Theorem 16. Let $(S,\overline{F})$ be a statistical metric space, U the two-parameter collection of subsets of $S \times S$ defined by

$$\mathcal{U} = U(\epsilon, \lambda);\ \epsilon > 0,\ \lambda > 0,$$

where

$$U(\epsilon, \lambda) = \left\{ (p,q); p,q \text{ in } S \text{ and } F_{pq}(\epsilon) > 1 - \lambda \right\},$$

and a two place function T from $[0,1] \times [0,1]$ to $[0,1]$ satisfying

$T(c,d) \geq T(a,b)$ for $c \geq a$, $d \geq b$ and $\sup_{x<1} T(x,x) = 1$. Suppose further that for all p,q,r in S and for all real numbers x,y, the Menger triangle inequality

$$F_{pr}(x + y) \geq T(F_{pq}(x), F_{qr}(y))$$

is satisfied. Then U is the basis for a Hausdorff uniformity on $S \times S$.

Before proving the theorem, one must first become familiar with the concept of diagonal, uniform space, and a basis for a uniform space.

To start with consider subsets of a Cartesian product $S \times S$ of a set S with itself. Such a subset is a relation U. A relation is a set of ordered pairs; and if U is a relation, the inverse relation U^{-1} is the set of all pairs (x,y) such that (y,x) are elements of U. The operation of taking inverses is involutory in the sense that $(U^{-1})^{-1}$ is always U. If $U = U^{-1}$, then U is called symmetric. If U and V are relations, then the composition $U \circ V$ is the set of all pairs (x,z) such that for some y it is true that (x,y) is an element of V and (y,z) is an element of U. The set of all pairs (x,x) for x in X is called the identity relation, or the diagonal and is denoted by Δ.

A uniformity for a set X is a non-void family $\mathcal{U}$ of subsets of $X \times X$ such that

(a) each member of $\mathcal{U}$ contains the diagonal Δ;

(b) if U is an element of $\mathcal{U}$, then U^{-1} is an element of U;

(c) if U is an element of $\mathcal{U}$, then $V \circ V$ is contained in $\mathcal{U}$ for some V in $\mathcal{U}$;

(d) if U and V are members of $\mathcal{U}$, then $U \cap V$ is an element of $\mathcal{U}$;

(e) if U is an element of $\mathcal{U}$ and $U \subset V \subset X \times X$, then V is contained in $\mathcal{U}$.

The pair $(X, \mathcal{U})$ is a uniform space.

The proof of Theorem 16 is an immediate consequence of the following theorem which is stated without proof.

Theorem 17. A family $\mathcal{B}$ of subsets of $X \times X$ is a base for some uniformity for X if and only if

(a) each member of $\mathcal{B}$ contains the diagonal Δ;

(b) if U is an element of $\mathcal{B}$, then U^{-1} contains a member of $\mathcal{B}$;

(c) if U is an element of $\mathcal{B}$, then $V \circ V \subset U$ for some V in $\mathcal{B}$;

(d) the intersection of two members of $\mathcal{B}$ contains a member of $\mathcal{B}$.

Proof of Theorem 16:

(a) Let $\Delta = \{(p,p);p \in S\}$ and $U(\epsilon,\lambda)$ be given. Since for any p an element of S, $F_{pp}(\epsilon) = 1$, it follows that (p,p) is an element $U(\epsilon,\lambda)$. Thus $\Delta \subset U(\epsilon,\lambda)$.

(b) Since $F_{pq} = F_{qp}$, $U(\epsilon,\lambda)$ is symmetric; hence $U(\lambda,\epsilon) = U^{-1}(\epsilon,\lambda)$.

(c) Let $U(\epsilon,\lambda)$ be given. One must show that there is a W, an element of $\mathcal{U}$, such that $W \circ W \subset U$. Choose $\epsilon' = \epsilon/2$ and λ' so small that $T(1 - \lambda',1 - \lambda') = 1 - \lambda$. Suppose that (p,q) and (q,r) belong to $W(\epsilon', \lambda')$, so that $F_{pq}(\epsilon') \geq 1 - \lambda'$ and $F_{qr}(\epsilon') > 1 - \lambda'$. Then

$$F_{pr}(x + y) \geq T(F_{pq}(x),F_{qr}(y))$$

$$F_{pr}(\epsilon) \geq T(F_{pq}(\epsilon'),F_{qr}(\epsilon')) \geq T(1 - \lambda', 1 - \lambda') \geq 1 - \lambda .$$

Thus, (p,r) is an element of $U(\epsilon,\lambda)$; this means that $W \circ W \subset U$.

(d) The intersection of $U(\epsilon,\lambda)$ and $U(\epsilon', \lambda')$ contains a member of $\mathcal{U}$, namely $U(\min(\epsilon, \epsilon'),\min(\lambda, \lambda'))$.

Thus, $\mathcal{U}$ is a basis for a uniformity on $S \times S$.

(e) If p and q are distinct, there exists an $\epsilon > 0$, such that $F_{pq}(\epsilon_0) = 1 - \lambda_0$. Consequently, (p,q) is not in $U(\epsilon_0, \lambda_0)$ and the uniformity generated by $\mathcal{U}$ is separated and, therefore, Hausdorff.

Note that the theorem is true in particular for all Menger spaces in which $\sup_{x<1} T(x,x) = 1$. However, it is true as well for many SM-spaces which are not Menger spaces.

<u>Corollary</u>. If $(S,\overline{F})$ is an SM-space satisfying the hypotheses of Theorem 16, then the sets of the form $N_p(\epsilon, \lambda) = \{q; \quad F_{pq} > 1 - \lambda\}$ are the neighborhood basis for a Hausdorff topology on S.

<u>Proof</u>: These sets are a neighborhood basis for the uniform topology on S derived from U.

<u>Theorem 18</u>. If an SM-space satisfies the hypotheses of Theorem 16, then the set determined by the sets $N_p(\epsilon, \lambda)$ is metrizable.

<u>Proof</u>: Let $\{(\epsilon_n, \lambda_n)\}$ be a sequence that converges to (0,0). Then this collection is a countable base. The metrization theorem found in [1] states that a uniform space is metrizable if and only if it is Hausdorff and its uniformity has a countable basis. In the previous work it was noted that a space satisfying the conditions of Theorem 16 was a Hausdorff space. It was also shown that the space defined in Theorem 16 was a uniformity. Now using these two results along with the countable basis, the conclusion of the theorem follows.

ACKNOWLEDGMENT

The writer wishes to express his gratitude and appreciation to Professor William L. Stamey, of the Department of Mathematics, for his help and consideration in formutation of this paper.

BIBLIOGRAPHY

(1) Kelly, John L.
General Topology, New York: D. Van Nostrand Company, Inc., 1955.

(2) Hall, Dick W. and Spencer, Guilford L.
Elementary Topology, New York: John Wiley and Sons, Inc., 1955.

(3) Menger, Karl
"Statistical Metrics," Proceedings National Academy of Science, U. S. A., Vol. 28, 1942, pp. 535-537.

(4) Schweizer, B. and Sklar, A.
"Statistical Metric Spaces," Pacific Journal Mathematics, Vol. 10, 1960, pp. 313-334.

(5) ________
"Topology and Tchebycheff," American Mathematical Monthly, Vol. 68, 1961, pp. 761-762.

(6) Schweizer, B., Sklar, A., and Thorp, E.
"The Metrization of Statistical Metric Spaces," Pacific Journal Mathematics, Vol. 10, 1960, pp.673-675.

(7) Thorp, Edward
"Best Possible Triangle Inequalities for Statistical Metric Spaces," Proceedings American Mathematical Society, Vol. 11, 1960, pp. 734-740.

(8) Wald, A.
"On a Statistical Generalization of Metric Spaces," Proceedings National Academy of Science, U. S. A., Vol. 29, 1943, pp. 196-197.

STATISTICAL METRIC SPACES

by

GARY ALLAN BOGAR

B. S., Missouri State Teachers College, 1960

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Mathematics

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1963

The concept of a statistical metric space was introduced by Karl Menger in 1942. In this paper he used distribution functions to define the distance between two points. Such a distribution function may be interpreted as the probability that the distance between p and q is less than x and is represented by $F_{pq}(x)$. The use of distribution functions results from the observation that the distance between two points is not given by a single measurement but instead is taken as the average of the distance obtained from several measurements.

The purpose of this paper is to investigate the properties of spaces defined by distribution functions and the topological properties of such spaces.

A statistical metric space is defined and four conditions are given, which it must satisfy. Upon examination of these four conditions, they are found to be similar to the four conditions for a metric space. In fact, it is proved that a metric space is a special case of a statistical space.

After defining a statistical metric space, the fourth condition (triangle inequality) is changed so as to define a Menger space. The Menger triangle inequality makes use of a special set of functions called t-functions. These t-functions are investigated and it is shown that given a t-function a Menger space is constructed for which it is the strongest, in a certain sense. Also for a special case of the t-function, the Menger space is found to be a statistical metric space.

Another change in the statistical metric triangle inequality gives a Wald space. The Wald space uses the convolution of two distribution functions to define the triangle inequality. Under a proper choice of the t-function, it is proved that a Wald space is equivalent to a Menger space. The Wald space is also found to be a statistical metric space where the Wald triangle inequality holds universally. Thus, the Menger and Wald spaces are both special cases of the statistical metric space.

The topology of a statistical metric space is induced by defining an ϵ, λ - neighborhood $N_p = \{p; F_{pq}(\epsilon) > 1 - \lambda\}$ for every $\epsilon > 0$, and $\lambda > 0$. The resulting space is found to be a Hausdorff space. By generalizing the above results, it is proved that a statistical metric space is metrizable. Therefore, in numerous instances the existence of a statistical metric implies the existence of an ordinary metric.