Czechoslovak Mathematical Journal

Karl Menger; Berthold Schweizer; Abe Sklar

On probabilistic metrics and numerical metrics with probability. I

Czechoslovak Mathematical Journal, Vol. 9 (1959), No. 3, 459–466

Persistent URL: http://dml.cz/dmlcz/100370

Чехословацкий математический журнал, т. 9 (84) 1959, Прага

ON PROBABILISTIC METRICS AND NUMERICAL METRICS WITH PROBABILITY 1

KARL MENGER, BERTHOLD SCHWEIZER, ABE SKLAR, Chicago

(Received November 11, 1958)

This note clarifies the relationship between the random metrics of A. Špaček and probabilistic metrics.

1. Probabilistic Metrics

Traditional metrics are numerical; that is to say, the distance $F(a, b)$ associated with any two points a, b, is a *number*. A probabilistic metric, as developed in [1, 2, 3, 4, 5, 7, 8],[1]) is defined by associating with a, b, not a number, but a cumulative distribution function F_{ab}, that is, a non-decreasing function, continuous to the left, and such that

$$\lim_{x\to\infty} F_{ab}(x) = 0 \quad \text{and} \quad \lim_{x\to\infty} F_{ab}(x) = 1 \, .$$

Certain general assumptions about the numerical distances in traditional metrics have as their counterparts simple assumptions about the distribution functions in probabilistic metrics. For each pair of points, a, b, the following postulates are made:

1. $F_{ab}(x) = 1$ for all positive x if and only if $a = b$.
2. $F_{ab}(0) = 0$.
3. $F_{ab}(x) = F_{ba}(x)$ for each x.

If $F_{ab}(x)$ is interpreted as the probability that the distance from a to b less than x, then Postulate 2 attributes the probability 0 to distances less than 0 and thus almost certainly rules out negative distances. According to Postulate 1, it is almost certain that the distance of each point from itself is less than any positive number. This, in conjunction with Postulate 2, makes it almost certain that the distance of each point from itself is 0. The distance from a point to a different point is positive with a positive probability. According to Postulate 3,

[1]) Numbers in brackets refer to the bibliography at the end of the paper.

for any x, the probability of the distance from a to b being less than x, and that of the distance from b to a being less than x, are equal.

A fourth postulate concerning triples of points (which corresponds to the traditional triangle inequality) need not be discussed here, since we will at first confine ourselves to the consideration of a space S consisting of exactly two points, p and q.

By a *probabilistic metric* on S we mean any ordered quadruple of distribution functions F_{pp}, F_{qq}, F_{pq}, F_{qp} satisfying Postulates 1—3.

A study which as yet has not been explicitly developed is that of *joint* distributions of distances. For instance, the probabilistic metric on the pair S might be based on a single 4-dimensional cumulative distribution function G whose value $G(u, v, x, y)$ might be interpreted as the probability of the following conjunction: the distances from p to p, from q to q, from p to q, and from q to p are $< u$, $< v$, $< x$, and $< y$, respectively. In this case F_{pp}, F_{qq}, F_{pq}, and F_{qp} would appear as the marginal distributions of G. In the papers quoted, however, it has been implicitly assumed that the distances are *independent*.[2]) This corresponds to the assumption

$$G(u, v, x, y) = F_{pp}(u) \, . \, F_{qq}(v) \, . \, F_{pq}(x) \, . \, F_{qp}(y) \quad \text{for all} \quad u, v, x, y \, . \tag{1}$$

As an application of probabilistic metrics one may, for instance, think of two fixed points and the distribution of their distances as observed in a large number of successive measurements.[3]) Postulates 1—3 reflect the facts that the distance between identical points is always found to be 0; that one never finds the distance between distinct points to be negative, but does find an appreciable proportion of positive distances; and that, for each $x > 0$, among a large number of measured distances from p to q, the proportion of results $< x$ is approximately equal to the proportion of results $< x$ among a large number of measured distances from q to p.

We call F_{ab} *rigid* at d (where d is a non-negative number) if

$$F_{ab}(x) = \begin{cases} 0 \text{ for } x \leqq d \, , \\ 1 \text{ for } x > d \, . \end{cases}$$

In any probabilistic metric, F_{pp} and F_{qq} are rigid at 0. If F_{pq} (and hence F_{qp}) is also rigid, then the probabilistic metric itself will be called *rigid*, more specifically, rigid with distance d.

Only when F_{pq} and F_{qp} are non-rigid can one speak of *proper* probabilistic metric. For only in that case do probabilities other than 0 and 1 occur and questions about probabilities of distances in the strict sense of the word probab-

[2]) This assumption of independence clearly underlies the triangle inequality as formulated by A. WALD [7], [8], where the convolution of the distributions of the two distances is taken as the distribution of their sum.

[3]) Various other applications are mentioned in [2].

ility arise.[4]) On the other hand, a rigid probabilistic metric (in which all distances are almost certain) clearly plays the same role as a traditional metric in the sense of Fréchet. Such a metric on S may be described by an ordered quadruple of numbers or by a point in the 4-dimensional Cartesian space C_4, namely, the point with coordinates

$$F(p, p), \; F(q, q), \; F(p, q), \; F(q, p)$$

(if $F(a, b)$ denotes the number at which the function F_{ab} is discontinuous), or, in still other words, by a function F whose domain consists of the four pairs

$$(\mathrm{S}^2) \qquad (p, p), \; (q, q), \; (p, q), \; (q, p)$$

and which satisfies the conditions

A. $F(p, p) = F(q, q) = 0.$

B. $F(p, q) > 0$ and $F(q, p) > 0.$

C. $F(p, q) = F(q, p).$

2. Numerical Metrics with Probability 1

A different connection between probabilistic and geometric concepts has recently been established by A. Špaček in his *Note on K. Menger's Probabilistic Geometry* [6].

We will call any real function defined on the class of all ordered pairs of points in a space an *infra-metric* on the space (the prefixes semi-, pseudo- and quasi- traditionally being used for other purposes). Instead of one probabilistic metric, Špaček has studied probabilities (i. e., σ-additive measure functions whose values belong to the interval $[0,1]$) defined on various classes of infra-metrics, and has sought to assign probabilities to the set of infra-metrics in such a manner that the class of all metrics has the probability 1. He calls a class of infra-metrics on which probabilities have been assigned in this manner a *random metric*.

Geometrically, in the case of a space S consisting of two points p and q, each infra-metric may be described by a point in C_4. Then Špaček's problem becomes that of setting up a probability measure π on subsets of C_4 in such a manner that the set $\mathfrak{M}$ corresponding to all metrics has the measure 1. Clearly, $\mathfrak{M}$ is the set of all points whose coordinates satisfy conditions A, B, C and, therefore, are of the form $(0, 0, d, d)$ for some $d > 0$. The set $\mathfrak{M}$ is a ray issuing from (but not including) the origin and bisecting the first quadrant in the plane of all

[4]) Normal distributions are ruled out by Postulate 2. Examples of possible distributions include: normal distributions that are truncated to the left at 0 or at any number > 0; exponential distributions; rectangular distributions; step functions with a finite or infinite number of points of discontinuity, etc.

points whose first two coordinates are 0. Every set $\mathfrak{F}$ that is disjoint from $\mathfrak{M}$ and for which $\pi(\mathfrak{F})$ is defined has the probability measure $\pi(\mathfrak{F}) = 0$.

The relation between Špaček's random metrics and probabilistic metrics becomes apparent when one translates, in the standard way, his assumptions about a probability measure, defined for certain sets of infra-metrics on S (or subsets of C_4). For any four numbers u, v, x, y, define $G(u, v, x, y)$ as the measure of the set of all infra-metrics F such that

$$F(p, p) < u\,, \quad F(q, q) < v\,, \quad F(p, q) < x\,, \quad F(q, p) < y\,;$$

or, in other words, let $G(u, v, x, y)$ be the measure of the set of all points in C_4 whose coordinates $F(p, p)$, $F(q, q)$, $F(p, q)$, $F(q, p)$ satisfy the above inequalities.

It should be noted that, conversely, given a probabilistic metric, one can, on the basis of assumptions about the joint distribution, introduce a measure on sets of infra-metrics. For instance, upon assuming independence, one can define $G(u, v, x, y)$ by formula (1). Any joint distribution function G induces a measure on C_4 and hence on infra-metrics.

The set $\mathfrak{M}$ is clearly the intersection of the sets $\mathfrak{F}_A, \mathfrak{F}_B, \mathfrak{F}_C$, of all infra-metrics satisfying Conditions A, B, C, respectively. In order that $\pi(\mathfrak{M}) = 1$ it is obviously necessary and sufficient [6] that

A′: $\pi(\mathfrak{F}_A) = 1$.

B′: $\pi(\mathfrak{F}_B) = 1$.

C′: $\pi(\mathfrak{F}_C) = 1$.

If, for each pair (a, b) belonging to (S^2) and for each number x, the set of all infra-metrics F such that $F(a, b) < x$ is denoted by $\mathfrak{F}^x_{ab}$, then the four functions F_{ab} defined by

$$F_{ab}(x) = \pi(\mathfrak{F}^x_{ab}) \text{ for each } x$$

are the marginal distributions of G. Špaček's Postulates A′ and B′ imply that the marginal distributions F_{pq} and F_{qp} satisfy[5]) Postulates 1 and 2. Moreover, taken together, Špaček's postulates imply that the entire probability is concentrated on the ray $\mathfrak{M}$. In order to see what this means we need two lemmas which we present without proof.

Lemma 1. If a 2-dimensional joint distribution function G defined on the xy-plane assigns the probability 1 to the line $y = x$, then the marginal distribution functions F_1 and F_2 defined by

$$F_1(x) = G(x, +\infty)\,, \quad F_2(y) = G(+\infty, y)$$

---

[5]) It should be noted that Condition B′ rules out distribution functions which have a positive limit at $0+$, for instance, normal distributions that are truncated to the left at 0.

are equal; that is to say, $F_1(t) = F_2(t)$ for all t. Furthermore, $G(x, y) = \text{Min}\,(F_1(x), F_2(y))$ for all x, y.

Lemma 2. If F_1 and F_2 are 1-dimensional distribution functions and $F_1(x) \cdot F_2(y) = \text{Min}\,(F_1(x), F_2(y))$, for all x, y, then at least one of the functions F_1 and F_2 is rigid.

Applying these lemmas to Špaček's conditions one obtains the following results.

Theorem 1. *If in a random metric in the sense of Špaček the distances from p to q and from q to p are independent, then the corresponding probabilistic metric is rigid.*

It follows that in such a random metric, for some number $d > 0$, $\pi(\mathfrak{F}) = 1$ or 0 for any measurable set $\mathfrak{F}$ according as $\mathfrak{F}$ does or does not include the point $(0, 0, d, d)$. Thus in the case of independent distances, a Špaček random metric concentrates the entire probability in one single point of C_4.

Parenthetically, we mention that, if F_{pq} is continuous and distances are independent, then not only is $\pi(\mathfrak{F}_C) < 1$, but $\pi(\mathfrak{F}_C) = 0$.

In a space containing more than two points, Theorem 1 of course applies to each pair of points.

Now, in a traditional metric on a space containing more than two points it is postulated that each triple p, q, r satisfies the triangle inequalities,

$$\text{D.}\quad F(p, q) + F(q, r) \geqq F(r, p)\,, \quad F(g, r) + F(r, p) \geqq F(p, q)\,,$$
$$F(r, p) + F(p, q) \geqq F(q, r)\,.$$

Correspondingly, in a random metric on a space containing more than two points, Špaček postulates the condition

$$\text{D}'.\quad \pi(\mathfrak{F}_D) = 1,$$

where $\mathfrak{F}_D$ is the set of all infra-metrics F satisfying D. From Theorem 1 or readily obtains

Theorem 2. *If in a random metric, the distances from a to b and from b to a are independent, for any two points a, b, then the resulting probabilistic metric is rigid and corresponds to a traditional metric in the sense of Fréchet.*

Coming back to a pair of points p, q and omitting the assumption that the two distances are independent, from the lemmas we infer the following:

Theorem 3. *To each random metric there corresponds a probabilistic metric in which the distances from p to q and from q to p are either independent and rigid or non-rigid and completely dependent.*[6]) *If a correlation exists it is* 1.

[6]) Observe that in both cases the distances are actually completely dependent. For, in the rigid case, the distance form p to q determines the distance from q to p and vice versa.

In a probabilistic metric corresponding to a random metric on a space of more than two points, the distances, say, from p to q and from p to r may or may not be independent. According to Theorems 2 and 3 it is sufficient to consider the case where, for each pair a, b, the distances from a to b and from b to a are completely dependent. If in this case, one further assumes that the lengths of the three sides of the triangle pqr are independent, then, in the presence of Conditions A′, B′, C′, imposing Condition D′ entails further severe restrictions on the distribution functions F_{pq}, F_{r}, F_{rp}. For instance:

1. Each of the three functions attains the value 1 for some finite argument.

2. At most one of the three functions has a point of increase[7]) at 0.

3. If 0 is a point of increase of one function, then the other two are rigid.

4. The sum of the largest points of increase of the three functions is not greater than twice the sum of their smallest points of increase.

As an example, consider a probabilistic triangle that is equilateral (i. e., for which F_{pq}, $F_{qr} = F_{rp}$) and for which the distribution functions are one and the same cumulative rectangular distribution, increasing from d_0 to d_1. Špaček's conditions imply that $d_1 \leqq 2d_0$.

3. Conclusions

The results of the preceding section show that probability measures on classes of infra-metrics in which the class of all metrics has probability 1 are extremely restrictive. In particular, in the case of independent distances, they rule out all *proper* probabilistic metrics. Indeed, as far as independent distances are concerned, in the last analysis, Špaček's random metrics simply reduce to traditional metrics, described in a round-about fashion.

In view of some remarks in Špaček's paper it must be strongly emphasized that the fact that the conditions for probabilistic metrics, as developed in the series of papers l. c.,[1]) are insufficient to ensure Špaček's conditions for a random metric, is by no means a weakness or a defect of these conditions. For even stronger conditions would be insufficient, as long as they admit proper probabilistic metrics with independent or not completely dependent distances, which in view of their possible applications to physics and psychometrics [1, 2], should certainly be admitted.

Even without the results of the preceding section, from general considerations it is evident that Špaček's concept of a random metric is too restrictive. For, while randomizing numerical metrics, he nevertheless places his main emphasis on those cases in which their occurrence will be certain — this despite

[7]) The distribution function F_{ab} is said to have a point of increase at c if $F_{ab}(c - \varepsilon) < F_{ab}(c + \varepsilon)$ for each $\varepsilon > 0$.

the fact that in the large realm of infra-metrics which he admits for comparison, the occurrence of a metric is an extremely unlikely event.

On the other hand, Špaček's ideas do lead to some interesting questions. Instead of concentrating, in the space of infra-metrics, the entire probability on metrics, we would suggest distributing the probability over a wider class of functions, which one might call near-metrics. For example, in the case of S, instead of concentrating the entire probability on the previously defined ray $\mathfrak{M}$, one might distribute the probability over a neighborhood of $\mathfrak{M}$, say a tube about $\mathfrak{M}$. A systematic study of such probability distributions, from the geometric point of view, would be a most useful outgrowth of Špaček's paper.

BIBLIOGRAPHY

[1] *K. Menger*: Statistical Metrics, Proc. Nat. Acad. Sci. USA, vol. *28* (1942), pp. 535—537.
[2] *K. Menger*: Probabilistic Geometry, ibid., vol. *37* (1951), pp. 226—229.
[3] *K. Menger*: Géométrie Générale. Memorial des Sci. Math., vol. *124*, Paris 1954.
[4] *B. Schweizer* and *A. Sklar*: Statistical Metric Spaces (forthcoming).
[5] *B. Schweizer* and *A. Sklar*: Espaces Métriques Aléatoires, Comptes Rendus Acad. Sci. Paris, vol. *247* (1958), pp. 2092—2094.
[6] *A. Špaček*: Note on K. Menger's Probabilistic Geometry, Czechoslovak Mathem. Journal, vol. *6* (1956), pp. 72—73.
[7] *A. Wald*: On a Statistical Generalization of Metric Spaces, Proc. Nat. Acad. Sci. USA, vol. *29* (1943), pp. 196—197.
[8] *A. Wald*: On a Statistical Generalization of Metric Spaces, Reports of a Math. Colloquium, 2nd. ser., vol. *5—6*, pp. 76—79.

Резюме

О ВЕРОЯТНОСТНЫХ МЕТРИКАХ И ЧИСЛЕННЫХ МЕТРИКАХ С ВЕРОЯТНОСТЬЮ 1

К. МЕНГЕР, Б. ШВЕЙЦЕР, Э. СКЛАР (K. Menger, B. Schweizer, A. Sklar), Chicago

(Поступило в редакцию 11/XI 1958)

Рассматриваются вероятностные метрики, которые были введены и исследованы в статьях [1]—[5], [7], [8], и случайные метрики, введенные в статье А. Шпачека [6]. Указывается, что всякая случайная метрика индуцирует естественным образом вероятностную метрику. Вероятностная метрика называется жесткой для точек p, q если расстояние от p до q имеет определенное значение с вероятностью 1; вероятностная метрика, жесткая для всех p, q называется жесткой; такая вероятностная метрика является, по существу, обычной численной метрикой.

Доказываются следующие результаты:

Теорема 1. *Если для случайной метрики (в смысле А. Шпачека) расстояния от p до q и от q до p независимы, то соответствующая вероятностная метрика является жесткой (для p, q).*

Теорема 2. *Если для случайной метрики при любых a, b расстояние от a до b и от b до a независимы, то индуцированная ею вероятностная метрика является жесткой и дает обычное метрическое пространство в смысле Фреше.*

Теорема 3. *Всякой случайной метрике соответствует вероятностная метрика, в которой расстояния от p до q и от q до p или независимы (так что метрика жестка для p, q) или полностью зависимы. Если существует их кореляционный коэффициент, то он равен* 1.

В связи с этими теоремами, обсуждаются соотношения между понятиями вероятностной и случайной (в смысле А. Шпачека) метрики и ограничения, вытекающие из наложенных в его определении условий. Указываются некоторые возможные модификации случайной метрики.