ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 1, Number 3, Summer 1971

A METRIC FOR WEAK CONVERGENCE OF DISTRIBUTION FUNCTIONS[1]

DAVID A. SIBLEY

Let Δ denote the set of distribution functions, that is, left-continuous nondecreasing functions from the real line into $[0, 1]$. The set of distribution functions of random variables (functions in Δ with sup 1 and inf 0) will be denoted by Δ_{rv}. The following facts are well known (see, e.g., [2]).

1. The space Δ is sequentially compact with respect to weak convergence. That is, any sequence of functions in Δ has a weakly convergent subsequence (Helly's First Theorem). However, Δ_{rv} does not have this property.

2. The set Δ is a metric space under the Lévy metric L defined for any F, G in Δ by

$$L(F, G) = \inf\{h;\ F(x - h) - h \leqq G(x) \leqq F(x + h) + h \text{ for all } x\}.$$

3. If (F_n) is a sequence in Δ_{rv} and F is also in Δ_{rv} then F_n converges weakly to F iff $L(F_n, F) \to 0$. Thus for sequences in Δ_{rv} whose limit is also in Δ_{rv} weak convergence and convergence in the L-metric are equivalent. The hypothesis that the limit belong to Δ_{rv} is necessary, for there are sequences in Δ_{rv} which converge weakly to a limit in Δ but do not converge in the Lévy metric. (One such sequence is discussed in this paper.)

Statement 3 shows that the relationship between weak convergence (in the sense of Helly's First Theorem) and convergence in the metric L is unsatisfactory. This state of affairs is due to the fact that the Lévy metric is sensitive to what happens at $+\infty$ and $-\infty$, while weak convergence is not. The purpose of this paper is to show that a modification of the Lévy metric yields a metric for Δ for which convergence corresponds precisely to weak convergence.

For any F, G in Δ and $h > 0$, define the properties

(1) $A(F, G; h)$ iff $F(x - h) - h \leqq G(x)$ for $-1/h < x < 1/h + h$,

(2) $B(F, G; h)$ iff $F(x + h) + h \geqq G(x)$ for $-h - 1/h < x < 1/h$,

and let

Received by the editors October 10, 1969.

AMS 1969 *subject classifications*. Primary 6020; Secondary 5435.

[1]Some of these results were part of the author's Senior Honors Thesis written under the direction of Professor Berthold Schweizer at the University of Massachusetts.

(3) $\mathcal{L}(F, G) = \inf\{h;$ both $A(F, G; h)$ and $B(F, G; h)$ hold$\}$.

Before showing that $\mathcal{L}$ is a metric with the desired properties, we make a few simple observations.

LEMMA 1. *$A(F, G; h)$ iff $B(G, F; h)$.*

PROOF. Substitute x for $x - h$ in (1).

LEMMA 2. *$A(F, G; h)$ is equivalent to*

$$F\left(x - \frac{h}{2}\right) - \frac{h}{2} \leqq G\left(x + \frac{h}{2}\right) + \frac{h}{2} \quad \text{for } -\left(\frac{h}{2} + \frac{1}{h}\right) < x < \left(\frac{h}{2} + \frac{1}{h}\right). \tag{4}$$

PROOF. Substitute $x + h/2$ for x in (1).

Finally, notice that $\mathcal{L}(F, G) \leqq 1$ for any F, G in Δ and that if $\mathcal{L}(F, G) = h > 0$, then both $A(F, G; h)$ and $B(F, G; h)$ hold.

THEOREM 1. *The function $\mathcal{L}$ defined by* (3) *is a metric for* Δ.

PROOF. (a) POSITIVITY. It is immediate that $\mathcal{L}(F, G) \geqq 0$ for F, G in Δ.

(b) IDENTITY. Clearly $\mathcal{L}(F, F) = 0$. Conversely, suppose $\mathcal{L}(F, G) = 0$. For any real x, we may choose h so small that both $A(F, G; h)$ and $B(F, G; h)$ hold on an interval containing x. Thus, we have

$$F(x - h) - h \leqq G(x) \leqq F(x + h) + h.$$

Letting $h \to 0$ yields $F(x -) \leqq G(x) \leqq F(x +)$. Hence, if x is a point of continuity of F, then $F(x) = G(x)$. It follows that $F = G$.

(c) SYMMETRY. This is an immediate consequence of Lemmas 1 and 2.

(d) TRIANGLE INEQUALITY. Let $\mathcal{L}(F, G) = h$ and $\mathcal{L}(G, H) = k$, where $h, k > 0$. For $h + k \geqq 1$, the inequality

$$\mathcal{L}(F, H) \leqq \mathcal{L}(F, G) + \mathcal{L}(G, H)$$

is trivial. Hence, we may assume that $h + k < 1$. For $a > 0$, let I_a denote the interval $[-a/2 - 1/a, \; a/2 + 1/a]$ and suppose $x \in I_{h+k}$. Using the fact that $h(h + k) < 1$, it is easy to show that $x - k/2 \in I_h$ and $x + h/2 \in I_k$. Then from $A(F, G; h)$ and $A(G, H; k)$, it follows that

$$\begin{aligned} F\left(x - \frac{h+k}{2}\right) - \frac{h+k}{2} &\leqq G\left(x - \frac{k}{2} + \frac{h}{2}\right) - \frac{k}{2} + \frac{h}{2} \\ &\leqq H\left(x + \frac{h+k}{2}\right) + \frac{h+k}{2}, \end{aligned}$$

whence $A(F, H; h + k)$ holds. Similarly, $B(F, G; h)$ and $B(G, H; k)$ imply $B(F, H; h + k)$. Thus, $\mathcal{L}(F, H) \leqq h + k = \mathcal{L}(F, G) + \mathcal{L}(G, H)$, completing the proof of the theorem.

The function $\mathcal{L}$ will be called the modified Lévy metric.

DEFINITION. A sequence (F_n) of functions in Δ converges weakly to a function F in Δ if $(F_n(x))$ converges to $F(x)$ at each point x of continuity of F. In this case, we write $F_n \xrightarrow{w} F$.

It is well known that $F_n \xrightarrow{w} F$ if and only if (F_n) converges to F pointwise on some dense subset of the real line [**1**]. Furthermore, since each F_n is in Δ, the limit F must be nondecreasing and may without loss of generality be assumed to be in Δ.

THEOREM 2. *Let (F_n) be a sequence of functions in Δ and let F be in Δ. Then $\mathcal{L}(F_n, F) \to 0$ as $n \to \infty$ if and only if $F_n \xrightarrow{w} F$.*

PROOF. First suppose $\mathcal{L}(F_n, F) \to 0$. Let x be a point of continuity of F and let $\epsilon > 0$ be given. Then for some $\delta > 0$, $|F(x) - F(y)| < \epsilon$ whenever $|x - y| < \delta$. Let $h = \min(\epsilon, \delta, 1/|x|)$ if $x \neq 0$, and $h = \min(\epsilon, \delta)$ if $x = 0$. Also let n be so large that $\mathcal{L}(F_n, F) < h$. Since $|x| < 1/h$, we have from (1) and (2) that

$$F_n(x) \geqq F(x - h) - h \geqq F(x) - 2\epsilon$$

and

$$F_n(x) \leqq F(x + h) + h \leqq F(x) + 2\epsilon,$$

whence $F_n(x) \to F(x)$ as $n \to \infty$. It follows that $F_n \xrightarrow{w} F$.

Conversely, suppose that $F_n \xrightarrow{w} F$. Then (F_n) converges to F pointwise on some set D which is dense in the real line. Let $\epsilon > 0$ be given, and choose a in D such that $a > \epsilon + 1/\epsilon$. Let $a_0 < -a$, a_0 in D, and partition the interval $[a_0, a]$ by choosing points $a_0 < a_1 < \cdots < a_m = a$, where a_i in D, $i = 0, 1, \cdots, m$, such that $|a_i - a_{i-1}| < \epsilon$ for $i = 1, 2, \cdots, m$. Choose N so large that for any $n \geqq N$ we have $|F_n(a_i) - F(a_i)| \leqq \epsilon/2$ for $i = 0, 1, \cdots, m$. Now, if either $-1/\epsilon \leqq x \leqq \epsilon + 1/\epsilon$ or $-\epsilon - 1/\epsilon \leqq x \leqq 1/\epsilon$, then x is in $[a_0, a]$. In both cases, there is k such that $a_{k-1} \leqq x \leqq a_k$. Consequently,

$$F_n(x) \leqq F_n(a_k) \leqq F(a_k) + \epsilon/2 \leqq F(x + \epsilon) + \epsilon/2,$$

and

$$F_n(x) \geqq F_n(a_{k-1}) \geqq F(a_{k-1}) - \epsilon/2 \geqq F(x - \epsilon) - \epsilon/2.$$

Thus, both $A(F_n, F; \epsilon)$ and $B(F_n, F; \epsilon)$ hold, whence $\mathcal{L}(F_n, F) < \epsilon$ and $\mathcal{L}(F_n, F) \to 0$ as $n \to \infty$.

An immediate consequence of Theorem 2 and Helly's First Theorem is

THEOREM 3. *The space* $(\Delta, \mathcal{L})$ *is compact.*

An obvious homotopy argument shows that any ϵ-neighborhood in $(\Delta, \mathcal{L})$ is pathwise connected, so that, in particular, $(\Delta, \mathcal{L})$ is both connected and locally connected. It follows from this and Theorem 3 that $(\Delta, \mathcal{L})$ is a Peano space, i.e., the continuous image of an arc.

The following example illustrates the difference between the metrics L and $\mathcal{L}$. Let θ denote the function which is identically zero and, for $n = 1, 2, \cdots$, define the step functions S_n by

$$(5) \qquad \begin{aligned} S_n(x) &= 0 \quad \text{if } x \leqq n, \\ &= 1 \quad \text{if } x > n. \end{aligned}$$

Then $S_n \xrightarrow{w} \theta$ and $\mathcal{L}(S_n, \theta) = 1/n$. On the other hand, $L(S_n, \theta) = 1$, for any n. In fact, $L(S_n, S_m) = \delta_{nm}$, so that (S_n) is not a Cauchy sequence in (Δ, L) and has no cluster point. Thus, in contrast to Theorem 3, (Δ, L) is not compact. But more can be said; for a similar construction shows that any L-open subset of Δ contains a sequence which converges weakly, but has no cluster point under the metric L. Thus, no compact subset in (Δ, L) can contain an L-open subset. Hence (Δ, L) is not even locally compact. Indeed, any compact subset of (Δ, L) is nowhere dense.

Finally, we note that if $F(x-h) - h \leqq G(x) \leqq F(x+h) + h$, for all real x, then both $A(F, G; h)$ and $B(F, G; h)$ hold. Thus, $\mathcal{L}(F, G) \leqq L(F, G)$, for any F, G in Δ. It follows that the $\mathcal{L}$-metric topology is weaker than the L-metric topology. The above example shows that in Δ we may replace "weaker" by "strictly weaker". However, in Δ_{rv} it can be shown that any L-open subset contains an $\mathcal{L}$-open set, so that L and $\mathcal{L}$ induce the same topology on Δ_{rv}.

The author is indebted to Professor Berthold Schweizer for his many valuable suggestions concerning this work.

REFERENCES

1. B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, GITTL, Moscow, 1949; English transl., Addison-Wesley, Reading, Mass., 1954. MR **12**, 839; MR **16**, 52.

2. M. Loève, *Probability theory*, 2nd ed., Van Nostrand, Princeton, N.J., 1960. MR **23** #A670.

CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIFORNIA 91109