A Metrization of Class-Convergences of Distributions

By Kinsaku TAKANO

(Received April 15, 1953)

1. Introduction

In the theory of probability it occurs very often, that for a given sequence of random variables $\{S_n\}$, there exist a sequence of positive numbers $\{a_n\}$ and a sequence of real numbers $\{b_n\}$, such that the sequence of the distributions of $(S_n-b_n)/a_n$ converges to some distribution. When only the existence of $\{a_n\}$ and $\{b_n\}$ are known, it is important to discuss the general method to determine the values of a_n and b_n. For this purpose, we have to consider the metrization of class-convergences of d.f.'s (distribution functions). It is well known that any arbitrary sequence of classes of d.f.'s $\{K_n\}$ converges to the improper class (A.I. Khintchine [2]). Therefore, we have to omit the improper class, when we consider the metrization of class-convergences. However, W. Doeblin [1] has shown that the convergence space of all proper classes is not metrizable. In this note, we shall discuss a metrization of the class-convergences except for an adequate neighbourhood of the improper class. It is sufficient to leave out the classes of d.f.'s $F(x)$ having the property that $\tilde{F}(+0)-\tilde{F}(-0)$ is near to 1, where $\tilde{F}(x)$ is the convolution of $F(x)$ and $[1-F(-x)]$:

$$\tilde{F}(x)=\int_{-\infty}^{\infty} F(x-y)d[1-F(-y)]\,.$$

Terminology: Two d.f.'s $F(x)$ and $G(x)$ belong to the same class, if and only if there exist a and b such that $F(ax+b)=G(x)$. In the convergence space of the d.f.'s, a sequence of d.f.'s $\{F_n(x)\}$ converges to a d.f. $F_0(x)$ ($\lim F_n(x)=F_0(x)$), if and only if $\{F_n(x)\}$ converges to $F_0(x)$ at every continuity point of the latter. A sequence of classes of d.f.'s $\{K_n\}$ converges to a class K_0, if and only if there exists a sequence of d.f.'s $\{F_n(x)\}$ such that $F_n \in K_n$ $(n=0, 1, 2,\dots)$ and $\lim F_n=F_0$. We denote P. Lévy's distance of two d.f.'s $F(x)$ and $G(x)$ by $d(F(x), G(x))$ or by $d(F, G)$ (P. Lévy [4], p. 47).

A sequence of d.f.'s $F_n(x)$ converges to a d.f. $F_0(x)$ if and only if $\lim_{n\to\infty} d(F_n, F_0)=0$.

2. A Metrization of Class-convergences of d.f.'s

Let $F(x)$ be a (one-dimensional) d.f., and let $\psi_F(l)$ be its m.c.f. (mean concentration function) introduced by K. Kunisawa [3]:

$$\psi_F(l)=\int_{-\infty}^{\infty}\frac{l^2}{l^2+x^2}d\tilde{F}(x), \qquad (0<l<\infty).$$

Clearly $\psi_F(l)$ is non-decreasing, continuous, and we have

$$0<\psi_F(l)\leq 1\,, \quad \psi_F(\infty)=1\,, \quad \psi_F(+0)=\tilde{F}(+0)-\tilde{F}(-0)\,.$$

The inverse of the m. c. f. $\psi_F(l)$,

$$D_F(\gamma)=\min\{l;\ \psi_F(l)\geq\gamma\}, \qquad (0<\gamma<1),$$

is called the dispersion function of the d.f. $F(x)$, and for any fixed γ, $(0<\gamma<1)$, we shall call $D_F(\gamma)$ the dispersion of $F(x)$ for parameter γ.

We shall use the following properties of the m. c. f.'s and the dispersion functions.

(1) $\quad \psi_F(+0)<1\,,\quad$ if $F(x)$ is proper.
$\quad\quad \psi_F(+0)=1\,,\quad$ if $F(x)$ is improper.

(2) If $\psi_F(+0)<\gamma$ then $D_F(\gamma)>0$, and conversely $(0<\gamma<1)$.

(3) $\quad \psi_{F(ax+b)}(l)=\psi_{F(x)}(al)\,,\quad D_{F(ax+b)}(\gamma)=\dfrac{1}{a}D_{F(x)}(\gamma)\,,$

$(a>0)$.

In case a sequence of d.f.'s $\{F_n(x)\}$ converges to a d.f. $F_0(x)$, it holds that

(4) $\quad \lim\psi_{F_n}(l)=\psi_{F_0}(l)\,,\qquad (0<l<\infty)\,,$

and

(5) $\quad \lim D_{F_n}(\gamma)=D_{F_0}(\gamma)\,,\qquad (0<\gamma<1)\,.$

(1)—(4) are easily proved and (5) is derived from (4) (see, for instance, K. Takano [5], Theorem 3). Both (4) and (5) hold at every point in the respective open interval, since neither $\psi_{F_0}(l)$ nor $D_{F_0}(\gamma)$ have discontinuity points. This is the reason why we use Kunisawa's m.c.f.'s instead of P. Lévy's maximal concentration functions.

From (3) we have

$$\psi_{F(ax+b)}(+0)=\psi_{F(x)}(+0)=\tilde{F}(+0)-\tilde{F}(-0)\,,\quad (a>0)\,.$$

Thus, $\psi_F(+0)$ is invariant as $F(x)$ runs in the same class K. We denote this value by $\psi_K(0)$. It is clear that

(6) $\quad 0\leq\psi_K(0)\leq 1, \qquad \psi_K(0)=1$ if and only if K is improper.

If $\psi_K(0)<\gamma<1$ (γ constant) and $F\in K$, then $D_F(\gamma)>0$ by (2). Write $F_1(x)=F(D_F(\gamma)x)$. Then from (3) we have

$$D_{F_1}(\gamma)=1\,,\qquad F_1\in K\,.$$

Thus, we obtain the following

Lemma 1. If a class K satifies $\psi_K(0) < \gamma < 1$ for a constant γ, there exists a d.f. $F(x)$ belonging to the class K such that $D_F(\gamma) = 1$.

Let Ω be the set of all proper classes, and for any fixed $\gamma(0 < \gamma < 1)$, let $\Omega_\gamma = \{K;\ \psi_K(0) < \gamma\}$. Then Ω_γ is monotone increasing with γ and $\Omega = \bigcup_\gamma \Omega_\gamma$.

On fixing γ, we shall define a metric in the set Ω_γ. Let K_1 and K_2 be two classes belonging to Ω_γ. From Lemma 1, we can choose two d.f.'s $F_1(x)$ and $F_2(x)$ such that $F_1 \in K_1$, $F_2 \in K_2$, $D_{F_1}(\gamma) = 1$ and $D_{F_2}(\gamma) = 1$. Any d.f. $F(x)$ such that $F \in K_1$ and $D_F(\gamma) = 1$, is equal to $F_1(x+c)$ for some constant c. Define the distance of the two classes K_1 and K_2 as follows:

$$d_\gamma(K_1, K_2) = \inf_{c_1, c_2} d(F_1(x+c_1), F_2(x+c_2)) = \inf_{-\infty<c<\infty} d(F_1(x), F_2(x+c)), \tag{7}$$

where $d(F_1(x), F_2(x))$ denotes P. Lévy's distance of two d.f.'s $F_1(x)$ and $F_2(x)$. Let $g(c) = d(F_1(x), F_2(x+c))$. Since $g(c)$ is continuous:

$$|g(c)-g(c')| \leqq d(F_2(x+c), F_2(x+c')) \leqq \sqrt{2}\,|c-c'|,$$

and

$$\lim_{c\to\pm\infty} g(c) = \sqrt{2} = \sup_c g(c),$$

it holds that

$$d_\gamma(K_1, K_2) = \min_{-\infty<c<\infty} d(F_1(x), F_2(x+c)) = \min_{-\infty<c<\infty} d(F_1(x+c), F_2(x)). \tag{8}$$

We have the following

THEOREM 1. *$d_\gamma(K_1, K_2)$ satisfies the axiom of distance in Ω_γ:*

$$d_\gamma(K_1, K_2) \geqq 0, \quad d_\gamma(K_1, K_2) = 0 \text{ if and only if } K_1 = K_2, \tag{9}$$

$$d_\gamma(K_1, K_2) = d_\gamma(K_2, K_1), \tag{10}$$

$$d_\gamma(K_1, K_3) \leqq d_\gamma(K_1, K_2) + d_\gamma(K_2, K_3). \tag{11}$$

Proof. (9) and (10) are evident. To prove (11), take a d.f. $F_2(x)$ from the class K_2 such that $D_{F_2}(\gamma)=1$. Then there exist $F_1(x)$ and $F_3(x)$ such that $F_1(x) \in K_1$, $F_3(x) \in K_3$, $D_{F_1}(\gamma)=1$, $D_{F_3}(\gamma)=1$ and $d_\gamma(K_1, K_2)=d(F_1, F_2)$ and $d_\gamma(K_2, K_3)=d(F_2, F_3)$. From the last two equalities, we have

$$d_\gamma(K_1, K_3) \leqq d(F_1, F_3) \leqq d(F_1, F_2) + d(F_2, F_3) = d_\gamma(K_1, K_2) + d_\gamma(K_2, K_3).$$

THEOREM 2. *Let $K_n \in \Omega_\gamma$ ($n=0, 1, 2, \ldots$). In order that the sequence of classes $\{K_n\}$ converges to the class K_0, it is necessary and sufficient that $\{K_n\}$ converges to K_0 in the sense of the above-defined distance.*

Proof. Sufficiency: Assume that $d_\gamma(K_n, K_0) \to 0$. Let $F_0(x) \in K_0$ be such that $D_{F_0}(\gamma)=1$. Then there exists $F_n(x) \in K_n$ such that $D_{F_n}(\gamma)=1$ and

$d_\gamma(K_n, K_0)=d(F_n, F_0)$ for every n. From this and the hypothesis, we have $\lim d(F_n, F_0)=0$ which means that $\{F_n(x)\}$ converges to $F_0(x)$ and, consequently, that $\{K_n\}$ converges to K_0.

Necessity: Assume that $\{K_n\}$ converges to K_0. Then there exists a sequence of d.f.'s $\{F_n(x)\}$ such that $F_n(x)\in K_n$ $(n=0,1,2,\dots)$ and

$$\lim_n F_n(x) = F_0(x)\,. \tag{12}$$

Let $D_n(\gamma)$ be the dispersion of $F_n(x)$ $(n=0,1,2,\dots)$. Then from (12) we have

$$\lim D_n(\gamma) = D_0(\gamma)\,, \tag{13}$$

and

$$D_n(\gamma) > 0\,, \quad (n=0,1,2,\dots\dots)\,, \tag{14}$$

making use of (5) and (2), respectively. From (12)—(14) we have

$$\lim F_n(D_n(\gamma)x) = F_0(D_0(\gamma)x)\,. \tag{15}$$

Since the dispersion of the d.f. $F_n(D_n(\gamma)x)$ for parameter γ is equal to 1, it holds that

$$d_\gamma(K_n,K_0)\leq d\Big(F_n(D_n(\gamma)x)\,,\ F_0(D_0(\gamma)x)\Big)\,. \tag{16}$$

As (15) implies that the right hand side of (16) tends to zero, we have $\lim d_\gamma(K_n, K_0)=0$ which is to be proved.

Making use of Theorem 2 we can give a simple proof to the following theorem of A. I. Khintchine [2].

Corollary: If a sequence of classes of d.f.'s converges to a proper class, then the limiting proper class is uniquely determined by the sequence.

Proof. Assume that a sequence $\{K_n\}$ has two limiting proper classes K_0 and K_0'. Then there exists γ $(0<\gamma<1)$ such that Ω_γ contains both K_0 and K_0'. Fix such a γ. For sufficiently large n, we have $K_n\in\Omega_\gamma$. Since the convergence space Ω_γ is metrizable, it holds that $K_0=K_0'$.

In the above we have used the fact that Ω_γ is open in the convergence space Ω. It is clear that Ω_γ is not closed in the convergence space Ω, however, we have the following

THEOREM 3. *The metric space Ω_γ is complete, i.e., any fundamental sequence converges.*

Proof. Assuming that $K_n\in\Omega_\gamma$ $(n=1,2,\dots)$ and that $\lim_{n,m\to\infty} d_\gamma(K_n, K_m)=0$, we shall prove that there exists a class K_0 such that $K_0\in\Omega_\gamma$ and $\lim_{n\to\infty} d_\gamma(K_n, K_0)=0$. From the assumption there exists a subsequence $\{K_{n(i)}\}$ such that $\sum_{i=1}^{\infty} d_\gamma(K_{n(i)}, K_{n(i+1)})<\infty$. Let $F_1(x)$ be a d.f. belonging to the class $K_{n(1)}$ such that $D_{F_1}(\gamma)=1$. Then we can choose successively $F_2(x)$, $F_3(x)$,...... such that $F_i(x)\in K_{n(i)}$, $D_{F_i}(\gamma)=1$ and $d(F_i, F_{i+1})=d_\gamma(K_{n(i)}, K_{n(i+1)})$, $(i=1,2,\dots)$. Since

$$d(F_p, F_q) \leq \sum_{i=p}^{q-1} d(F_i, F_{i+1}) = \sum_{i=p}^{q-1} d(K_{n(i)}, K_{n(i+1)}), \quad (p < q),$$

it holds that $\lim_{p,q\to\infty} d(F_p, F_q)=0$. From the completeness of the metric space of d.f.'s with P. Lévy's distance, there exists a d.f. $F_0(x)$ such that $\lim_{i\to\infty} d(F_i, F_0) =0$, from which we have $\lim_{i\to\infty} D_{F_i}(\gamma) = D_{F_0}(\gamma)$. As $D_{F_i}(\gamma) = 1$ we have $D_{F_0}(\gamma) = 1$ which implies that $\psi_{F_0}(+0) < \gamma$. Let K_0 be the class containing the d.f. $F_0(x)$. Then K_0 belongs to Ω_γ and it holds that $\lim_{i\to\infty} d_\gamma(K_{n(i)}, K_0)=0$. From the assumption we have $\lim_{n,n(i)\to\infty} d_\gamma(K_n, K_{n(i)})=0$. Therefore it holds that $\lim_{n\to\infty} d_\gamma(K_n, K_0) = 0$.

For any two classes K_1 and K_2 belonging to Ω_γ, there exist d.f.'s $F_1(x)$ and $F_2(x)$ such that $F_1(x) \in K_1$, $F_2(x) \in K_2$, $D_{F_1}(\gamma) = 1$, $D_{F_2}(\gamma) = 1$ and $d_\gamma(K_1, K_2) = d(F_1, F_2)$. In this case we may assume that one of the two d.f.'s has zero as a median, but it is not always possible to choose them so that both have zero as medians.

Example: Let

$$F_1(x) = \begin{cases} 0, & x < -a, \\ (x+a)/2a, & -a < x < a, \\ 1, & x > a, \end{cases}$$

$$F_2(x) = \begin{cases} 0, & x < -1, \\ (x+1)/2, & -1 < x < 0, \\ 1, & x > 0. \end{cases}$$

Then we have

$$\psi_{F_1}(+0) = 0, \qquad \psi_{F_2}(+0) = 1/4.$$

For some γ $(1/4 < \gamma < 1)$, it holds that $D_{F_2}(\gamma) = 1$. Fixing such γ we can determine a so that $D_{F_1}(\gamma)=1$. It is easy to verify that $1/2 < a < 1$, using

$$\psi_{F_1}(l) = \frac{l}{a}\tan^{-1}\frac{2a}{l} - \frac{l^2}{4a^2}\log\left(1+\frac{4a^2}{l^2}\right)$$

and

$$\psi_{F_2}(l) = \frac{1}{4} + l\tan^{-1}\frac{1}{l} - \frac{l^2}{4}\log\left(1+\frac{1}{l^2}\right).$$

Let K_1 and K_2 be the classes which contain $F_1(x)$ and $F_2(x)$, respectively. Then it holds that

$$d_\gamma(K_1, K_2) = d(F_1(x+a/2), F_2(x)) < d(F_1(x), F_2(x)).$$

The medians of $F_1(x)$ and $F_2(x)$ are both determined uniquely and both equal to zero.

We notice that if the median of a d. f. $F(x)$ is uniquely determined, so is also the median of a d. f. $F(ax+b)$ $(a > 0)$.

THEOREM 4. *Let $\{K_n;\ n = 0, 1, 2,\dots\}$ be a sequence in Ω_T, and assume that any d. f. belonging to K_0 has the uniquely determined median. For any n, let $F_n(x) \in K_n$ be any d. f. with median 0 and dispersion (for parameter γ) 1. Then, in order that $\{K_n\}$ converges to K_0, it is necessary and sufficient that $\{F_n(x)\}$ converges to $F_0(x)$.*

Proof of Necessity. Assume that $\{K_n\}$ converges to K_0. Then there exist sequences $\{a_n\}$ and $\{b_n\}$ such that

$$\text{(17)} \qquad \lim F_n(a_n x + b_n) = F_0(a_0 x + b_0)\,, \qquad (a_n > 0,\ n = 0, 1, 2, \dots).$$

As $F_n(a_n x + b_n)$ has the dispersion $1/a_n$ and median $-b_n/a_n$ for any n $(n = 0, 1, 2, \dots)$, from (17) we have

$$\text{(18)} \qquad \begin{cases} \lim (1/a_n) = 1/a_0\,, & \text{and} \\ \lim (-b_n/a_n) = -b_0/a_0\,, & \end{cases}$$

(see, for instance, K. Takano [5], Corollary 2 to Theorem 3) which has the result that $a_n \to a_0$ and $b_n \to b_0$. Therofore, from (17), it holds that $\lim F_n(x) = F_0(x)$.

Using this Theorem 4, we can solve the problem stated in the introduction. However, it seems natural to do so by using (5) and the following

Lemma 2. Let $\{F_n(x);\ n = 1, 2, \dots\}$ be a sequence of d.f.'s, let $G(x)$ be a d. f., let $\{a_n\}$ and $\{\alpha_n\}$ be sequences of positive numbers, and let $\{b_n\}$ and $\{\beta_n\}$ be sequences of real numbers. Assume that

$$\lim F_n(a_n x + b_n) = G(x)\,,$$
$$\lim \frac{\alpha_n}{a_n} = 1\,, \qquad \lim \frac{\beta_n - b_n}{a_n} = 0\,.$$

Then we have

$$\lim F_n(\alpha_n x + \beta_n) = G(x)\,.$$

(K. Takano [5], Corollary 3 to Theorem 4.)

We shall denote the disperson of the d.f. of a random variable X by $D_X(\gamma)$.

THEOREM 5. *Let $\{S_n;\ n = 1, 2, \dots\}$ be a sequence of random variables and $G(x)$ be a d. f. Assume that $D_G(\gamma) > 0$ and $D_{S_n}(\gamma) > 0$, $(n = 1, 2, \dots)$, for some γ, $(0 < \gamma < 1)$. Fixing such a γ, write $D = D_G(\gamma)$ and $D_n = D_{S_n}(\gamma)$. Then in order that there exists a sequence of positive numbers $\{a_n\}$ such that the sequence of the d. f.'s of S_n/a_n converges to $G(x)$, it is necessary and sufficient that the sequence of the d. f.'s of*

$$\frac{S_n}{D_n/D}$$

converges to $G(x)$.

THEOREM 6. *Assume moreover that the median of* $G(x)$ *is uniquely determined and is equal to 0. Let* $\{m_n\}$ *be a sequence of medians corresponding to* $\{S_n\}$. *Then, in order that there exist a sequence of positive numbers* $\{a_n\}$ *and a sequence of real numbers* $\{b_n\}$ *such that the sequence of the d.f.'s of* $(S_n-b_n)/a_n$ *converges to* $G(x)$, *it is necessary and sufficient that the sequence of the d.f.'s of*

$$\frac{S_n-m_n}{D_n/D}$$

converges to $G(x)$.

3. Case when P. Lévy's dispersion is used

Let $F(x)$ be a d.f. and $Q_F(l)$ and $L_F(\gamma)$ be respectively its maximal concentration function and dispersion function introduced by P. Lévy (P. Lévy [4], p. 44). $Q_F(+0)=\sup_x[F(x+0)-F(x-0)]$ is invariant as $F(x)$ runs in the same class K. We denote this value by $Q_K(0)$. Substitute $Q_F(l)$, $L_F(\gamma)$ and $Q_K(0)$ for $\psi_F(l)$, $D_F(\gamma)$ and $\psi_K(0)$, respectively, in the preceding argument, then Theorem 1 holds and Theorem 2 and 4 hold if the dispersion function of a d.f. belonging to K_0 is continuous at γ.

INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] W. DOEBLIN: Sur l'ensemble de puissances d'une loi de probabilité, *Studia Math.*, 1940.

[2] A. I. KHINTCHINE: Ueber Klassenkonvergenz von Verteilungsgesetzen, *Mitt. Forrsch. Inst. Math. u. Mech. Univ. Tomsk*, 1, 1937.

[3] K. KUNISAWA: On an analytical method in the theory of independent random variables, *Ann. Inst. Stat. Math.*, Vol. 1, No. 1, 1949.

[4] P. LÉVY: *Théorie de l'addition des variables aléatoires*, 1937.

[5] K. Takano: On the convergence of classes of distributions, *Ann. Inst. Stat. Math.*, Vol. 3, No. 1, 1951.