PLISKA
STUDIA MATHEMATICA BULGARICA

ПЛИСКА
БЪЛГАРСКИ МАТЕМАТИЧЕСКИ СТУДИИ

HAUSDORFF METRIC CONSTRUCTION IN THE PROBABILITY MEASURES SPACE

SVETLOZAR T. RACHEV

The aim of this paper is to carry out an analysis of the Lévy and Lévy— Prohorov metrics as well of the uniform metrics and of other metrics in the space of probability measures. This is achieved by comparing their metric structures with the structure of the Hausdorff distance.

Introduction. In the solution of a number of problems of probability theory the method of metric distance functions has successfully been used for a long time (Lévy [9], Kolmogorov [8], Prohorov [11], Esséen [5] Strassen [20], Dudley [4], Zolotarev [23], Cambanis, Simons, Stout [2]). The essence of this method is based on the knowledge of properties o metrics in the space of random variables as well as on the principle according tf, which in every problem of the approximating type a metric as a comparisoo measure must be selected in accordance with the requirements to its propern ties (Zolotarev [23]).

In this paper properties of Hausdorff metric structure are used to solve two well-known problems considered earlier by other authors.

The first problem deals with the analysis of the "minimal" property of Lévy, Lévy — Prohorov and uniform metrics, as well of the other metrics (Strassen [20], Dudley [4], Zolotarev [23]).

The second problem represents a generalization of the Lévy and Lévy — Prohorov metric structure (Varadarajan [22], Zolotarev [23]).

The results of this paper are essentially contained in Rachev [12, 13, 14, 15].

1. Probability metrics and their properties. We would like to start by briefly mentioning the definitions and some properties of probability metrics (for general acquaintance we recommend papers Zolotarev [23, 24, 25]).

Denote by $\mathfrak{X}=\{X\}$ a set of random variables defined in some probability space $(\Omega, \mathscr{A}, \mathsf{P})$ and taking values in a certain separable metric space (U, d) and let $(\mathfrak{X})_k$ be a space of joint distributions of all possible sets $(X_1, X_2, \ldots, X_k)$ of random variables from $\mathfrak{X}$.

In the space $\mathfrak{X}$ the mapping $\mu:(\mathfrak{X})_2\to[0, \infty]$ is called a probability metric (or simply a metric) in the case when it possesses the metric properties of "symmetry", "triangle inequality" and the following analogue of the "identification" property (Zolotarev [23]):

$$\mathsf{P}(X=Y)=1 \Rightarrow \mu(X, Y)=0. \tag{1.1}$$

The metric $\mu(X, Y)$ is called simple if its value is completely determined by the pair of marginal distributions $\mathscr{L}(X)$, $\mathscr{L}(Y)$ and a compound metric in all remaining cases. In case of a simple metric (1.1) is equivalent to the condition $\mathscr{L}(X)=\mathscr{L}(Y)\Rightarrow\mu(X, Y)=0$. If we require also the converse, i. e. $\mu(X, Y)=0\Rightarrow\mathscr{L}(X)=\mathscr{L}(Y)$ one obtains the usual concept of a metric, but only in the space of marginal distribution $(\mathfrak{X})_1$. In th is case we can use both forms $\mu(X, Y)$ and $\mu(\mathscr{L}(X), \mathscr{L}(Y))$.

We give some examples of simple and compound metric in $\mathfrak{X}$. The Ky Fan compound metric (distance in probability):

$$\mathscr{K}(X, Y)=\inf\{\varepsilon>0: \mathsf{P}(d(X, Y)>\varepsilon)\leqq\varepsilon\}.$$

The indicator compound metric $i(X, Y)=\mathsf{E}I\{X\neq Y\}$ (I is the indicator function)˙ The Lévy—Prohorov simple metric $\pi(X, Y)=\inf\{\varepsilon>0: \mathsf{P}(X\in A)\leqq\mathsf{P}(Y\in A^\varepsilon)+\varepsilon$' $A\in\mathfrak{B}$, where $\mathfrak{B}$ is the system of all Borel sets in (U, d) and $A^\varepsilon=\{u; d(u, A)\leqq\varepsilon\}$. The simple distance in variation

$$\sigma(X, Y)=\sup\{|\mathsf{P}(X\in A)-\mathsf{P}(Y\in A)|.$$

Every metric μ in $\mathfrak{X}$ is related to the so-called minimal metric $\widehat{\mu}(X, Y)=\inf\mu(X, Y)$, where the infimum is taken over the set of all possible joint distributions $\mathscr{L}(X, Y)\in(\mathfrak{X})_2$ of the random variables X, Y with fixed marginal distribution $\mathscr{L}(X)$ and $\mathscr{L}(Y)$ (Zolotarev [23]).

The relationship between $\mathscr{K}$ and π was established in the well-known paper Strassen [20] (see also Dudley [4]). Strassen proved that

$$\widehat{\mathscr{K}}=\pi. \tag{1.2}$$

Dobrushin [3] established a similar relationship:

$$\widehat{i}=\sigma. \tag{1.3}$$

Recently, the notion of a minimal metric turned out to be a very usefull one in such problems as continuity and stability of stochastic models (Zolotarev [23 — 26]). In particular the inequality of the type $\nu(X, Y)\leqq\psi(\mu(X, Y)$ (the conditions on ψ being quite general) implies analogous inequality between the minimal metrics (Zolotarev [23]).

$$\widehat{\nu}(X, Y)\leqq\psi(\widehat{\mu}(X, Y)). \tag{1.4}$$

Let (S, ρ) be a metric space with metric ρ, and $\mathfrak{A}(S)$ be the set of all nonempty subsets of S. A Hausdorff distance (Hausdorff [6]) between two elements of $\mathfrak{A}(S)$ is defined as

$$r(G_1, G_2)=\max\{\sup_{x_1\in G_1}\inf_{x_2\in G_2}\rho(x_1, x_2), \sup_{x_2\in G_2}\inf_{x_1\in G_1}\rho(x_1, x_2)\}. \tag{1.5}$$

Sendov and Penkov [16] defined the Hausdorff distance between two bounded functions on the real line R. They noticed also that for distribution functions the Lévy distance coincides with the Hausdorff distance.

2. Minimal metrics in the random variables space

2.1. Lévy metric and uniform distance in the distribution function space. Let us denote by $\mathfrak{F}$ — the set of distribution functions on the real line R. The Lévy metric in the space $\mathfrak{F}$ metrizes the weak topology.

Let as define, for every $\lambda>0$, the Lévy metric as follows:

$$L_\lambda(F, G)=\inf\{\varepsilon>0: G(x-\lambda\varepsilon)-\varepsilon\leqq F(x)\leqq G(x+\lambda\varepsilon)+\varepsilon, \text{ for all } x\in R\}. \tag{2.1}$$

Properties. 1) For every $\lambda>0$, $L_\lambda(F, G)$ is a metric in $\mathfrak{F}$ and $L_1=L$, where L is the usual Lévy metric;

2) On the interval $(0, \infty)$, $L_\lambda(F, G)$ is a nonincreasing functions of λ;
3) Take the uniform distance $\rho(F, G)$ in $\mathfrak{F}$.

$$\rho(F, G)=\sup\{|F(x)-G(x)|, \quad x\in R\}$$

and $W(F, G)=\sup\{|F^{-1}(t)-G^{-1}(t)|, \ t\in[0, 1]\}$, where $F^{-1}(t)=\sup\{x: F(x)\leq t\}$. Then we have

$$\lim_{\lambda\to 0} L_\lambda(F, G)=\rho(F, G), \tag{2.2}$$

$$\lim_{\lambda\to\infty} \lambda L_\lambda(F, G)=W(F, G). \tag{2.3}$$

The following Theorems 1 and 1* give some Hausdorff distance representation of the Lévy metric.

Theorem 1. *For every* $\lambda>0$

$$L_\lambda(F, G)=\max\{\sup_{x_1\in R}\inf_{x_2\in R}\max\{\frac{1}{\lambda}|x_1-x_2|, F(x_2)-G(x_1)\}, \tag{2.4}$$

$$\sup_{x_2\in R}\inf_{x_1\in R}\max\{\frac{1}{\lambda}|x_1-x_2|, \quad G(x_1)-F(x_2)\}\}.$$

In accordance to Sendov [17, 18] let us define the Hausdorff distance between the bounded functions on R.

Suppose that for $\lambda>0$, $\rho_\lambda(A_1, A_2)$, $A_1=(x_1, y_1)$, $A_2=(x_2, y_2)$ is the Minkowski distance in the plane:

$$\rho_\lambda(A_1, A_2)=\max\{\frac{1}{\lambda}|x_1-x_2|, \ |y_1-y_2|\}. \tag{2.5}$$

If G_1, G_2 are closed nonempty sets in the plane, then the Hausdorff distance $r_\lambda(G_1, G_2)$ according to (1.5), is equal to

$$r_\lambda(G_1, G_2)=\max\{\max_{A_1\in G_1}\min_{A_2\in G_2}\rho_\lambda(A_1, A_2), \max_{A_2\in G_2}\min_{A_1\in G_1}\rho_\lambda(A_1, A_2)\}.$$

Denote be F_R the set of all closed point sets on the plane which are bounded and convex in relation of the x-axis and whose projections on the y-axis coincide with R. Let $f(x)$ be a bounded function on R. The intersection $\bar{f}$ of all set of F_R, containing the graph of f is called its complete graph. For every $\lambda>0$, let us define the Hausdorff distance between their complete graphs, i. e $r_\lambda(f, g)=r_\lambda(\bar{f}, \bar{g})$.

Theorem 1* (Sendov, Penkov). *For every* $\lambda>0$ *and* $F, G\in\mathfrak{F}$

$$L_\lambda(F, G)=r_\lambda(F, G).$$

The proof is similar to the proof contained in Rachev [12] for the case $\lambda=1$.

Now, let us define in the space of random variables the following functional K_λ: For every $\lambda>0$

$$K_\lambda(X, Y)=\inf\{\varepsilon>0: \mathsf{P}(X<x-\lambda\varepsilon, \ Y\geqq x)\leqq\varepsilon,$$
$$\mathsf{P}(Y<x-\lambda\varepsilon, \ X\geqq x)\leqq\varepsilon, \text{ for all } x\in R\}.$$

This functional has the following properties:

1) For every $\lambda>0$, $K_\lambda(X, Y)$ is a compound metric in $\mathfrak{X}$;

2) On the interval $(0, \infty)$, $K_\lambda(X, Y)$ is a nonincreasing function of λ;

3) Denote by $\Delta(X, Y)$, $w(X, Y)$ the following metrics in $\mathfrak{X}$:

$$\Delta(X, Y)=\max\{\sup_{x\in R} \mathsf{P}(X<x\leq Y), \sup_{x\in R} \mathsf{P}(Y<x\leq X)\},$$

$$w(X, Y)=\inf\{\varepsilon>0;\ \mathsf{P}(X<x-\varepsilon, Y\geq x)=0,$$
$$\mathsf{P}(Y<x-\varepsilon, X\geq x), \text{ for all } x\in R\}.$$

Then the relations

$$\lim_{\lambda\to 0} K_\lambda(X, Y)=\Delta(X, Y), \quad \lim_{\lambda\to\infty} \lambda K_\lambda(X, Y)=w(X, Y)$$

are true;

4) The metrics K_λ, Δ and w are weakly regular, i. e. for every triplet of random variables X, Y, Z such that the pare X, Y and the random variable Z are independent one has $K_\lambda(X+Z, Y+Z)\leq K_\lambda(X, Y)$;

5) For every $\lambda>0$ and for all $c\neq 0$ we have: $K_\lambda(cX, cY)=K_{\lambda/|c|}(X, Y)$, $\Delta(cX, cY)=\Delta(X, Y)$, $w(cX, cY)=|c|w(X, Y)$, i. e. from 4) follows that K_λ is a perfect (1, 0) metric; Δ is an ideal metric of order zero, while w is of order one. (The definitions of the notions perfect and ideal metrics are given by Zolotarev [23];

Denote by $a\vee b=\max(a, b)$.

6) For every $\lambda>0$

$$L_\lambda(X, Y)\vee L_\lambda(X, X\vee Y)\vee L_\lambda(Y, X\vee Y)\leq K_\lambda(X, Y), \tag{2.6}$$

$$\Delta(X, Y)=\rho(X, X\vee Y)\vee\rho(Y, X\vee Y). \tag{2.7}$$

From property (2.6) follows that $K_\lambda(X_n, Y)\to 0$ implies a weak convergence $L(X_n, Y)\to 0$, $L(X_n\vee Y, Y)\to 0$, when $n\to\infty$. The equality (2.7) shows that the Δ — convergence is equivalent to a uniform convergence $\rho(X_n, Y)\to 0 \rho(X_n\vee Y, Y, Y)\to 0$. From (2.6) and (2.7) follows that if $F_Y(x)=\mathsf{P}(Y<x)$ is a continuous distribution function, then $K(X_n, Y)\to 0$ if and only if $\Delta(X_n, Y)\to 0$;

7) For every $\lambda>0$

$$K_\lambda(X, Y)=\max\{\sup_{x_1\in R}\inf_{x_2\in R}\max\{\tfrac{1}{\lambda}|x_1-x_2|, \mathsf{P}(X\geq x_1, Y<x_2)\}, \tag{2.8}$$

$$\sup_{x_2\in R}\inf_{x_1\in R}\max\{\tfrac{1}{\lambda}|x_1-x_2|, \mathsf{P}(X<x_1, Y\geq x_2)\}\}.$$

Theorem 2. *For every $\lambda>0$ the equality $\widehat{K}_\lambda=L_\lambda$ holds.*

Corollary. $\widehat{\Delta}(X, Y)=\rho(X, Y)$, $\widehat{w}(X, Y)=W(X, Y)$.

Let $\mathfrak{A}=\{a\}$ be a linear space of real functions of bounded variation, defined on whole axis x. We shall consider in the space $\mathfrak{A}$ some norm Λ which apart from obligatory properties possesses the following one: if $0\leq a_1(x)\leq a_2(x)$, $x\in R$, then $\Lambda(a_1)\leq\Lambda(a_2)$.

Consider the probability metrics $\nu(X, Y)=\Lambda(|F_X-F_Y|)$, where $F_X(x)=\mathsf{P}(X<x)$ and $\mu(X, Y)=\Lambda(\mathsf{P}(X<x\leq Y)+\mathsf{P}(Y<x\leq X))$.

Theorem 3. *For every* $X, Y \in \mathfrak{X}$ *the equality* $\nu(X, Y) = \widehat{\mu}(X, Y)$ *holds.*

The functionals $\Lambda_p(a) = \{\int_{-\infty}^{+\infty} |a(x)|^p dx\}^{1/p}, p \geq 1, \Lambda_\infty(a) = \sup(|a(x)|, x \in R\}$ may be used as examples of the norms Λ. In the case $\Lambda = \Lambda_1$ Theorem 3 is in fact, the Kantorovich—Rubinstein statement $\widehat{E}|X-Y| = \int_{-\infty}^{+\infty} |F_X(x) - F_Y(x)| dx$ (Kantorovich and Rubinstein [7], Vallander [21], Cambanis et al. [2], Dudley [4]).

2.2. The metric H in the distribution functions space. In this part we shall consider the metric H_λ, $\lambda \in (0, \infty)$, in the space $\mathfrak{F}$, which are topologically stronger than the Lévy metric L_λ.

For every $\lambda > 0$ we define the following functional in $\mathfrak{F}$

$$(2.9) \qquad H_\lambda(F, G) = \max\{\sup_{x_1 \in R} \inf_{x_2 \in R} \max\{\frac{1}{\lambda}|x_1 - x_2|, |F(x_1) - G(x_2)|\},$$

$$\sup_{x_2 \in R} \inf_{x_1 \in R} \max\{\frac{1}{\lambda}|x_1 - x_2|, |F(x_1) - G(x_2)\}\}.$$

The functional $H_\lambda(F, G)$ has the following properties:

1. For every $\lambda > 0$, $H_\lambda(F, G)$ is a metric in $\mathfrak{F}$.

Introduce the metric $h_\lambda(B_1, B_2)$, $B_1, B_2 \subset D = R \times [0, 1]$ as a Hausdorff distance, induced by ρ_λ (2.5). Define as a closure graph $\Gamma(F)$ the following subset of the space D

$$\Gamma(F) = [\bigcup_{x \in R}(x, F(x+0))] \cup [\bigcup_{x \in R}(x, F(x-0))].$$

Then

2. $H_\lambda(F, G) = h_\lambda(\Gamma(F), \Gamma(G))$.

3. If F and G are continuous distribution functions, then $H_\lambda = L_\lambda$.

4. For every $\lambda > 0$, $L_\lambda \leq H_\lambda \leq \lim_{\lambda \to 0} H_\lambda = \rho$.

Note. Let two metrics d_1 and d_2 be given in the metric space S. Then the metric d_1 is called stonger than d_2, $d_1 \succ d_2$ if the d_1 — convergence of the sepuence of elements in S implies d_2 — convergence, and if the opposite statement is not true;

5. For every $\lambda > 0$ the relation $L_\lambda \prec H_\lambda \prec \rho$ holds.

We describe now the topological conditions for convergence in the metric H_λ, using the fact that the metric H_λ is Hausdorff distance in $D = R \times [0, 1]$.

Let us denote by $\operatorname{lt} B_n$ the topological limit of the sequence of the sets $B_n \subset D$, $n = 1, 2, \ldots$ (Hausdorff [6]).

Theorem 4. *If* F, F_n *are distribution functions, then* $H_\lambda(F_n, F) \to 0$ *if and only if* $\operatorname{lt}\Gamma(F_n)$ *exists and coincides with* $\Gamma(F)$.

The metric space $\mathfrak{F}$ with metric H is separable, but noncomplete. We shall define a metric d_0 satisfying the following two conditions:

(i) d_0 is topologically equivalent to the metric H;

(ii) the metric space $\mathfrak{F}$ with the metric d_0 is complete. The metric d_0 will have the Skorohod metric structure (Skorohod [19], Billingsly [1]). Let $\mathfrak{Z}$ be a space of strongly increasing continuous functions $\lambda(t)$ such that $\lambda(-\infty) = -\infty$, $\lambda(+\infty) = +\infty$. Consider in the space $\mathfrak{Z}$ the following functionals

$$\|\lambda\|_1 = \sup_{x_1 \neq x_2} |\log \frac{\lambda(x_1) - \lambda(x_2)}{x_1 - x_2}|, \quad \|\lambda\|_2 = \sup_{x \in R} |\lambda(x) - x|.$$

Then for every $F, G \in \mathfrak{F}$ we define

$$d_0(F, G) = \inf_{\lambda \in \mathscr{L}} \max \{ \|\lambda\|_1, \|\lambda\|_2, \sup_{x \in R} |F(x) - G(\lambda(x))| \}.$$

For every $A \subset \mathfrak{F}$ and $\delta > 0$ denote

$$\tilde{\omega}_A(\delta) = \sup \min \{ \sup_{A \in \mathfrak{F}} [F(t) - F(t_1)], \sup_{A \in \mathfrak{F}} [F(t_1 + \delta) - F(t)] \},$$

where the supremum is taken over t and t_1, such that $t_1 \leq t \leq t_1 + \delta$. For $A = \{F\}$ we denote $\tilde{\omega}_F = \tilde{\omega}_A$.

Note. A subset A of a metric space (S, d) is said to be d-compact if every sequence of points in A has a subsequence d-converging to a point of S. Some authors refer to this property as relative sequential compactness.

Theorem 5. *The set A of distribution functions is H-compact if and only if A satisfies the following conditions:*
(A) *A — is weakly compact (L-compact),*
(B) $\lim_{\delta \to 0} \sup_{F \in A} \tilde{\omega}_F(\delta) = 0.$

Corollary. *The set* $A \subset \mathfrak{F}$ is ρ-*compact iff and only iff A is weakly compact and* $\lim_{\delta \to 0} \tilde{\omega}_A(\delta) = 0$.

We shall define a compound metric T_λ in such a way that H_λ should be minimal with respect to T_λ. For every $\lambda > 0$ denote

$$T_\lambda(X, Y) = \max \{ \sup_{x_1 \in R} \inf_{x_2 \in R} \max \{ \tfrac{1}{\lambda} |x_1 - x_2|, \mathsf{P}(X \geq x_1, Y < x_2), \mathsf{P}(Y \geq x_2, X < x_1) \}$$

$$\sup_{x_2 \in R} \inf_{x_1 \in R} \max \{ \tfrac{1}{\lambda} |x_1 - x_2|, \mathsf{P}(X \geq x_1, Y < x_2), \mathsf{P}(Y \geq x_2, X < x_1) \} \}.$$

The functional $T_\lambda(X, Y)$ has the following properties:

1) For every $\lambda > 0$, $T_\lambda(X, Y)$ is a probability metric in $\mathfrak{X}$;
2) For every $\lambda > 0$, $K_\lambda \leq T_\lambda \leq \lim_{\lambda \to 0} T_\lambda(X, Y) = \Delta(X, Y)$.

Theorem 6. *For every* $\lambda > 0$, $\hat{T}_\lambda = H_\lambda$.

3. Weighted Lévy metric in the distribution functions space. The weighted metrics are used in many problems of estimation of the remainder term in the central limit theorem (Petrov [10]).

Let $q(x) \geq \alpha > 0$, $x \in R$ be a continuous function on the real line. When using the Lévy metric structure (2.1) to determine the weighted Lévy metric, i. e.

$$L^*(F, G; q) = \inf \{ \varepsilon > 0 \colon q(x) [F(x) - G(x + \varepsilon) \leq \varepsilon, \quad q(x) [G(x) - F(x + \varepsilon)] \leq \varepsilon, \text{ for all } x \in R \} \tag{3.1}$$

we need to suppose that $q(x)$ is a nondecreasing function. However, by keeping the Hausdorff structure for the Lévy metric (2.4) we can use it without this restriction.

Let $\alpha = q(x_0) = \inf \{ q(x), x \in R \}$ and $q_1(x) = q(x) I\{x \leq x_0\} + \alpha I\{x \geq x_0\}$, $q_2(x) = \alpha I\{x \leq x_0\} + q(x) I\{x \geq x_0\}$. Let us define the weighted Lévy metric as follows:

$$(3.2)\quad L(F, G; q)=\max\{\sup_{x_1\in R}\inf_{x_2\in R}\max[|x_1-x_2|, q_1(x_2)F(x_2)-q_1(x_1).G(x_1)],$$

$$\sup_{x_2\in R}\inf_{x_1\in R}\max[|x_1-x_2|, q_1(x_1)G(x_1)-q_1(x_2)F(x_2)],$$

$$\sup_{x_1\in R}\inf_{x_2\in R}\max[|x_1-x_2|, q_2(x_1)\bar{F}(x_1)-q_2(x_2)\bar{G}(x_2)],$$

$$\sup_{x_2\in R}\inf_{x_1\in R}\max[|x_1-x_2|, q_2(x_2).\bar{G}(x_2)-q_2(x_1)\bar{F}(x_1)]\},$$

where $\bar{F}=1-F$.

$L(F, G; q)$ is a simple probability metric and $L(F, G; 1)=L(F, G)$. For every continuous function $q(x)\geqq\alpha>0$ and $\varepsilon>0$ denote by $A_q(x; \varepsilon)$, $x\in R$, the following function

$$A_q(x; \varepsilon)=\sup\{q(x)/q(y); y\in[x-\varepsilon, x+\varepsilon]\}.$$

Let Q be a space of continuous functions satisfying the following condition $\overline{\lim}_{\varepsilon\to 0}\overline{\lim}_{|x|\to\infty} A_q(x; \varepsilon)<+\infty$.

The functions $\alpha+\exp\{|x|^\gamma\}$, $\gamma\in(0, 1]$ and $\alpha+|x|^\beta$, $\beta>0$ may be used as examples of the functions $q\in Q$, but $\alpha+\exp\{|x|^\gamma\}\notin Q$, when $\gamma>1$.

Theorem 7. *If $q\in Q$, then $\mathfrak{F}$ is a complete metric space with metric $L(F, G; q)$.*

Let $\mathfrak{F}_q$ be the set of all $F\in\mathfrak{F}$ such that

$$\lim_{N\to\infty}\sup_{x<-N} q(x)F(x)=0 \quad\text{and}\quad \lim_{N\to\infty}\sup_{x>N} q(x)\bar{F}(x)=0.$$

Theorem 8. *Let $q\in Q$. Then the set A of distribution functions is a $L(\cdot, \cdot, q)$ — compact subset of $\mathfrak{F}_q$ if and only if* $\lim_{N\to\infty}\sup_{x\in A}\sup_{x<-N} q(x)\times F(x)=0$ *and* $\lim_{N\to\infty}\sup_{x\in A}\sup_{x>N} q(x)\bar{F}(x)=0$.

Define, for every $F\in\mathfrak{F}$ the following functional

$$M_q(F)=\max\{\sup_{x\leqq 0} q(x)F(x), \sup_{x\geqq 0} q(x)\bar{F}(x)\}.$$

Theorem 9. *Let $q\in Q$. If $F\in\mathfrak{F}$ and $F_n\in\mathfrak{F}_q$, $n=1, 2, \ldots$ then $L(F_n, F; q)\to 0$ if and only if $L(F_n, F)\to 0$ and $M_q(F_n)\to M_q(F)$.*

4. The Lévy metric in random vector spaces. Denote by R^k — k-dimentional Euclidean space with norm $|x|=\max_{1\leqq i\leqq k}|x_i|$ and by $\mathfrak{F}^k$ — the set of distribution functions on the R^k. Let e be the unit vector in R^k. The Lévy metric $L_\lambda(F, G)=\inf\{\varepsilon>0; G(x-\lambda\varepsilon e)-\varepsilon\leqq F(x)\leqq G(x+\lambda\varepsilon e)+\varepsilon$, for every $x\in R^k\}$, $\lambda>0$, metricizes the weak topology in $\mathfrak{F}^k$.

Denote by $\mathfrak{F}(F_1, \ldots, F_k)$ the set of all distributions $F\in\mathfrak{F}^k$ with one-dimensional marginal distributions $F_1, F_2, \ldots, F_k$.

Theorem 10. *For every $\lambda>0$*

$$\min\{L_\lambda(F, G); F\in\mathfrak{F}(F_1, \ldots, F_k)\ G\in\mathfrak{F}(G_1, \ldots, G_k)\}=\max_{1\leqq i\leqq k} L_\lambda(F_i, G_i).$$

Denote by $\rho(F, G)$ the uniform distance in $\mathfrak{F}^k$:

$$\rho(F, G)=\sup\{|F(x)-G(x)|, x\in R^k\}.$$

Corollary. $\min\{\rho(F, G), F\in\mathfrak{F}(F_1, \ldots, F_k), G\in\mathfrak{F}(G_1, \ldots, G_k)\}=\max\{\rho(F_i, G_i), i=1, \ldots, k\}$.

The Lévy metric has the following Hausdorff distance representation: For every $\lambda>0$

$$L_\lambda(F, G)=\max\{\sup_{x_1\in R^k}\inf_{x_2\in R^k}\max\{\frac{1}{\lambda}\|x_1-x_2\|, G(x_2)-F(x_1)\},$$

$$\sup_{x_2\in R^k}\inf_{x_1\in R^k}\max\{\frac{1}{\lambda}\|x_1-x_2\|, F(x_1)-G(x_2)\}\}.$$

5. Hausdorff metric consrtruction as a generalization of the Lévy — Prohorov distance structure

5.1. Hausdorff-distance representation of Lévy — Prohorov metric. Let $\mathscr{C}$ be the set of all closed subsets on the metric space (U, d). The Hausdorff distance $r(G_1, G_2)$ metricizes $\mathscr{C}$. Let $\mathscr{P}(\mathfrak{B})$ be the system of all probability measures on $\mathfrak{B}=\mathfrak{B}_d$. If $\mathsf{P}\in\mathscr{P}$ then $\mathscr{C}$determines $\mathscr{R}$ uniquely.

Definition (Prohorov [11]). *Let* P, Q *be two Borel prabability measures on a metric space* (U, d). *Let* $\pi(\mathsf{P}, \mathsf{Q})=\inf\{\varepsilon>0;\ \mathsf{P}(C)\leq\mathsf{Q}(C^\varepsilon)+\varepsilon,\ \mathsf{Q}(C)\leq\mathsf{P}(C^\varepsilon)+\varepsilon$, for all $C\in\mathscr{C}\}$, *where* $C^\varepsilon=\{x\in U;\ d(x, C)\leq\varepsilon\}$.

Let us define, for every $\lambda>0$, *the Lévy — Prohorov metric as follows:*

(5.1) $\pi_\lambda(\mathsf{P}, \mathsf{Q})=\inf\{\varepsilon>0:\ \mathsf{P}(C)\leq\mathsf{Q}(C^{\lambda\varepsilon})+\varepsilon,\ \mathsf{Q}(C)\leq\mathsf{P}(C^{\lambda\varepsilon})+\varepsilon,$ *for* all $C\in\mathscr{C}\}$.

The Lévy — Prohorov metric has the following properties:

1) For every $\lambda>0$, $\pi_\lambda(\mathsf{P}, \mathsf{Q})$ is a metric in $\mathscr{P}$.

Denote the uniform metric on sets

$$\sigma(\mathsf{P}, \mathsf{Q})=\sup\{|\mathsf{P}(A)-\mathsf{Q}(A)|,\quad A\in\mathfrak{B}\},$$

$$e(\mathsf{P}, \mathsf{Q})=\inf\{\varepsilon>0;\quad \mathsf{P}(A)\leq\mathsf{Q}(A^\varepsilon),\ \text{for all } A\in\mathfrak{B}\}.$$

2) From (5.1) it follows

$$\lim_{\lambda\to 0}\pi_\lambda(\mathsf{P}, \mathsf{Q})=\sigma(\mathsf{P}, \mathsf{Q}), \tag{5.2}$$

$$\lim_{\lambda\to\infty}\lambda\pi_\lambda(\mathsf{P}, \mathsf{Q})=l(\mathsf{P}, \mathsf{Q}). \tag{5.3}$$

The Lévy — Prohorov metric has the following Hausdorff-distance representation:

Theorem 11. *For every* $\lambda>0$

$$\pi_\lambda(\mathsf{P}, \mathsf{Q})=\max\{\sup_{C_1\in\mathscr{C}}\inf_{C_2\in\mathscr{C}}\max\{\frac{1}{\lambda}r(C_1, C_2), \mathsf{P}(C_1)-\mathsf{Q}(C_2)\}, \tag{5.4}$$

$$\sup_{C_2\in\mathscr{C}}\inf_{C_1\in\mathscr{C}}\max\{\frac{1}{\lambda}r(C_1, C_2), \mathsf{Q}(C_2)-\mathsf{P}(C_1)\}\}.$$

Let $(\Omega, \mathscr{A}, \mathsf{P})$ be a probability space, $\mathscr{X}=\{X\}$ the set of random variables on $(\Omega, \mathscr{A}, \mathsf{P})$ taking values in a separable metric space (U, d). Let us define, for every $\lambda>0$ the Ky Fan metric

$$\mathscr{K}_\lambda(X, Y)=\inf\{\varepsilon>0:\ \mathsf{P}(d(X, Y)>\lambda\varepsilon)\leq\varepsilon\}. \tag{5.5}$$

Denote the indicator metric $i(X, Y)=\mathsf{E}I\{X\neq Y\}$ and $\mathscr{L}(X, Y)=\operatorname{ess\,sup} d(X, Y)$ $=\inf\{\varepsilon>0:\ \mathsf{P}(d(X, Y)>\varepsilon)=0\}$.

Then we have

$$\lim_{\lambda\to 0} \mathscr{K}_\lambda(X, Y) = i(X, Y), \tag{5.6}$$

$$\lim_{\lambda\to\infty} \lambda\mathscr{K}_\lambda(X, Y) = \mathscr{L}(X, Y). \tag{5.7}$$

Combining the relation $\widehat{\mathscr{K}} = \pi$, see (1.2), with (1.4) and (5.2), (5.3), (5.6), (5.7) we show that

$$\widehat{\mathscr{K}}_\lambda(X, Y) = \pi_\lambda(X, Y), \tag{5.8}$$

$$\widehat{i}(X, Y) = \sigma(X, Y), \tag{5.9}$$

$$\widehat{\mathscr{L}}(X, Y) = l(X, Y). \tag{5.10}$$

Another proof of (5.9) is given by Dobrushin [3]. The relation (5.10) gives an answer of a Dudley's question [4, n. 20. 1].

Further on we consider compound probability metric β_λ, $\lambda>0$ in the space $\mathfrak{X}$ which have Hausdorff metric structure and $\widehat{\beta}_\lambda = \pi_\lambda$. For every $\lambda>0$ we define the following functional in $\mathfrak{X}$:

$$\beta_\lambda(X, Y) = \inf\{\varepsilon>0: \mathsf{P}(X\in C,\ Y\notin C^{\lambda\varepsilon}) \leqq \varepsilon,\ \mathsf{P}(Y\in C,\ X\notin C^{\lambda\varepsilon}) \leqq \varepsilon, \text{ for all } C\in\mathscr{C}\}.$$

The functional $\beta_\lambda(X, Y)$ has the following properties:

1. For every $\lambda>0$, $\beta_\lambda(X, Y)$ is a metric in $\mathfrak{X}$.

Denote the following compound metric in $\mathfrak{X}$

$$B(X, Y) = \max\{\sup_{C\in\mathscr{C}} \mathsf{P}(X\in C,\ Y\notin C),\ \sup_{C\in\mathscr{C}} \mathsf{P}(X\notin C,\ Y\in C)\},$$

$$C(X, Y) = \inf\{\varepsilon>0: \mathsf{P}(X\in C,\ Y\notin C^\varepsilon) = 0,$$

$$\mathsf{P}(Y\in C,\ X\notin C^\varepsilon) = 0, \text{ for all } C\in\mathscr{C}\}.$$

2. We have $\lim_{\lambda\to 0} \beta_\lambda(X, Y) = B(X, Y)$, $\lim_{\lambda\to\infty} \lambda\beta_\lambda(X, Y) = C(X, Y)$. For any set $A\subset U$ let ∂A denote the boundary of A.

3. If $\beta_\lambda(X_n, X)\to 0$ then

$$\mathsf{P}(X_n\in A)\to\mathsf{P}(X\in A),\ \mathsf{P}(X_n\in A,\ X\in A)\to\mathsf{P}(X\in A)$$

for all A with $\mathsf{P}(X\in\partial A) = 0$. $B(X_n, X)\to 0$ if and only if $\sigma(X_n, X)\to 0$ and

$$\sup\{|\mathsf{P}(X_n\in A,\ X\in A) - \mathsf{P}(X\in A)|,\ A\in\mathfrak{B}\}\to 0.$$

4. For every $\lambda>0$ we have $\widehat{\beta}_\lambda = \pi_\lambda$, $\widehat{B} = \sigma$, $\widehat{C} = l$.

5. For $\lambda>0$

$$\beta_\lambda(X, Y) = \max\{\sup_{C_1\in\mathscr{C}} \inf_{C_2\in\mathscr{C}} \max\{\frac{1}{\lambda} r(C_1, C_2),\ \mathsf{P}(X\in C_1,\ Y\notin C_2)\},$$

$$\sup_{C_2\in\mathscr{C}} \inf_{C_1\in\mathscr{C}} \max\{\frac{1}{\lambda} r(C_1, C_2),\ \mathsf{P}(Y\in C_2,\ X\notin C_1)\}\}.$$

5.2. The Lévy — Prohorov distance in the space of continuous from above functions on sets. Let (U, d) be a metric space and $\mathscr{P}$ be a set of

Borel probability measures on U. Let the single-valued, real function $\varphi(C)$ be defined in the metric space $(\mathscr{C}, r)$ of all close sets $C \subset U$ with Hausdorff metric r. Denote $\mathscr{E} = \mathscr{E}(a, b)$ — the set of all functions $\varphi : \mathbf{A}(\varphi) \to [a, b]$, $\mathbf{A}(\varphi) \subset \mathscr{C}$, which are continuous from above, i. e. if $C_n \in \mathbf{A}(\varphi)$ and $C \in \mathbf{A}(\varphi)$, and $r(C_n, C) \to 0$ then $\lim \sup_n \varphi(C_n) \leq \varphi(C)$.

Note. If $\mathsf{P} \in \mathscr{P}$, then $\mathsf{P} \in \mathscr{E}(0, 1)$.

Definition. *Let $\varphi_k : \mathbf{A}_k \to [a, b]$ be a sequence of continuous from above functions. We call the function*

$$\bar{\varphi}(C) = \sup \{\overline{\lim_{k'}}\, \varphi_{k'}(C_{k'}),\ r(C_{k'}, C) \xrightarrow[k' \to \infty]{},\ C_{k'} \in \mathbf{A}_{k'},\ \{\mathbf{A}_{k'}\}_{k'} C\{\mathbf{A}_k\}_k\},$$

$$\bar{\varphi}(C) : \bar{\mathbf{A}} \to [a, b]$$

an upper topological limit $\bar{\varphi} = \overline{\mathrm{lt}}\, \varphi_k$ of the sequence $\varphi_{k'}$. The function $\bar{\varphi}$ is defined in $\bar{\mathbf{A}}_k = \overline{\mathrm{lt}}\, \mathbf{A}_k$ — the upper closed limit of set $\mathbf{A}_k$ (Hausdorff [6]).

Definition. *The function*

$$\underline{\varphi}(C) = \sup \{\underline{\lim_{k}}\, \varphi_k(C_k),\ r(C_k, C) \to 0,\ C_k \in \mathbf{A}_k,\ k = 1, 2, \ldots\}$$

is called a lower topological limit $\underline{\varphi} = \underline{\mathrm{lt}}\, \varphi_k$. The function $\underline{\varphi}$ is defined in $\underline{\mathbf{A}} = \underline{\mathrm{lt}}\, \mathbf{A}_k$ the lower topological limit of sets $\mathbf{A}_k$. If $\underline{\mathbf{A}} = \bar{\mathbf{A}} = \mathrm{lt}\, \mathbf{A}_k$ and for any $C \in \overline{\mathrm{lt}}\, \bar{\mathbf{A}}_k$

(5.11) $$\underline{\varphi}(C) = \bar{\varphi}(C)$$

then the function $\varphi = \mathrm{lt}\varphi_k = \underline{\mathrm{lt}}\, \varphi_k = \overline{\mathrm{lt}}\, \varphi_k$ is called a topological limit of $\{\varphi_k\}$.

Denote by $\mathsf{P}_n \Rightarrow$ the weak convergence of probability measures in $\mathscr{P}$.

Theorem 12. *If P, P_n, $n = 1, 2, \ldots$ are Borel probability measures, then $\mathsf{P} = \mathrm{lt}\, \mathsf{P}_n$ if and only if $\mathsf{P}_n \Rightarrow \mathsf{P}$.*

Definition. *Let $\varphi_i : \mathbf{A}_i \to [a, b]$, $i = 1, 2$ be two continuous from above functions. Define the Levy—Prohorov metric between φ_1 and φ_2 by*

(5.12) $$\pi(\varphi_1, \varphi_2) = \max \{\sup_{C_1 \in \mathbf{A}_1} \inf_{C_2 \in \mathbf{A}_2} \max\{r(C_1, C_2),\ \varphi_1(C_1) - \varphi_2(C_2)\},$$
$$\sup_{C_2 \in \mathbf{A}_2} \inf_{C_1 \in \mathbf{A}_1} \max\{r(C_1, C_2)\ \varphi_2(C_2) - \varphi_1(C_1)\}\}.$$

If $\varphi_1, \varphi_2 \in \mathscr{P}$ then the representation (5.12) *in accordance with* (2.4) *coincides with the usual Lévy—Prohorov metric.*

Denote by $\mathscr{E}_0$ the space of all $\varphi \in \mathscr{E}$ with closed definition ranges.

Definition. *If $\varphi_n \in \mathscr{E}$, $\varphi_0 \in \mathscr{E}$ and $\pi(\varphi_n, \varphi_0) \to 0$ then the sequence $\{\varphi_n\}$ is called metrically convergent and $\varphi_0 = \lim \varphi_n$ its metric limit.*

Theorem 13. *If the metric limit $\varphi = \mathrm{lm}\, \varphi_k$ exists, then $\varphi = \mathrm{lt}\, \varphi_k$.*

Theorem 14. *If the sequence $\{\varphi_0\} \subset \mathscr{E}$ is a π-fundamental one, then $\overline{\mathrm{lt}}\, \varphi_k = \underline{\mathrm{lt}}\, \varphi_k$.*

Note that in the space $\mathscr{P}$ the implication

$$\pi(\mathsf{P}_n, \mathsf{P}_m) \xrightarrow[n \to \infty,\ m \to \infty]{} 0 \text{ leads } \exists\ \mathsf{P} \in \mathscr{P} : \mathsf{P}_n \underset{n \to \infty}{\Rightarrow} \mathsf{P}$$

fails for non-Polish space. From theorem 14 it follows that it $\pi(P_n, P_m) \to 0$ when $n, m \to \infty$ then the limit $\varphi = \mathrm{lt}\, P_n \in \mathscr{E}$ exists.

Theorem 15. *If (U, d) is a complete metric space, then the metric spece $(\mathscr{E}_0, \pi)$ is also complete.*

Theorem 16. *If (U, d) is compact, then the metric space $(\mathscr{E}_0, \pi)$ is also compact.*

Acknowledgement. I should like to express my thanks to V. M. Zolotarev for discussing the results and important comments.

REFERENCES

1. P. Billingsley. Convergence of Probability Measures. New York, 1968.
2. S. Cambanis, G. Simons, W. Stout. Inequalities for $Ek(X, Y)$ when the marginals are fixed. *Z. Wahrscheilichkeitstheorie verw. Geb.*, **36**, 1976, 284—294.
3. Р. Л. Добрушин. Задание системы случайных величин при помощи условий распределений. *Теория вероятн. примен.*, **15**, 1970, 469—497.
4. R. M. Dudley. Probability and metrics. *Aarhus Univ., Lect. Notes Ser.*, No 45, 1976.
5. C-G. Esséen. On the mean central limit theorems. *Trans. Royal Inst. Technol.* Stockholm, **21**, 1958, 1—31.
6. F. Hausdorff. Set Theory. 3rd ed. New York, 1962.
7. Л. Канторович, Р. Рубинштейн. Об одном функциональном пространстве и некоторых экстремальных задачах. *Доклады АН СССР*, **115**, 1957, 1058—1061.
8. А. Н. Колмогоров. Некоторые работы последних лет в области предельных теорем теории вероятностей. *Вестн. Моск. унив.*, **10**, 1953, 29—38.
9. P. Lévy. L'arithemétique des lois de probabilités. *J. Math. Pures Appl.*, **103**, 1953, 17—40.
10. В. В. Петров. Суммы независимых случайных величин. Москва, 1972.
11. Ю. В. Прохоров. Сходимость случайных процессов и предельные теоремы теории вероятностей. *Теория вероятн. примен.*, **1**, 1956, 177—238.
12. С. Т. Рачев. Конструкция Хаусдорфа в пространстве вероятностных мер. *Зап. научн. сем. ЛОМИ АН СССР*, **87**, 1979, 87—103.
13. С. Т. Рачев. Расстояние Леви—Прохорова в пространстве полунепрерывных функций множеств. — В: Труды семинара: Устойчивость стохастических моделей. Москва, **1980**, **76—88**.
14. С. Т. Рачев. Минимальные метрики в пространстве вещественных случайных величин *Доклады АН СССР*, **257**, 1981, 1061—1070.
15. S. T. Rachev. Minimal metrics in the real valued random variables space. *Lecture Notes in Math.*, **982**, 1983, 136—150.
16. Б. Сендов, Б. Пенков. ε-ентропия и ε-капацитет на пространството от непрекъснатите функции. *Изв. Мат. Инст. БАН*, **6**, 1962, 27—50.
17. Б. Сендов. Хаусдорфовые приближения. София, 1979.
18. Б. Сендов. Некоторые вопросы теории приближений функций и множеств в хаусдорфовой метрике. *Успехи мат. наук.*, **24**, 1969, 141—173.
19. А. В. Скороход. Предельные теоремы для случайных процессов. *Теория вероятн. примен.*, **1**, 1956, 261—290.
20. V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, **36**. 1965, 423—439.
21. С. С. Валландер. Вычисление расстояния по Вассерштейну между распределениями вероятностей на прямой. *Теория вероятн. примен.*, **18**, 1973, 824—827.
22. В. Варадарайн. Меры на топологических пространствах. *Мат. сборник*, **55**, 1961 35—100.
23. В. М. Золотарев. Математические расстояния в пространствах случайных величин и их распределений. *Мат. сборник*, **101**, 1976, 104—141.
24. В. М. Золотарев. Идеальные метрики в задачах аппроксимации распределения сумм независимых случайных величин. *Теория вероятн. примен.*, **22**, 1976, 749—765.
25. V. M. Zolotarev. General problems of the stability of mathematical models. *Bull. Int. Statist. Inst.*, **47**, 1977, 382—401.
26. V. M. Zolotarev. Ideal metrics in the problems of probability theory and mathematical statistics. *Austral. J. Statist.*, **21**, 1979, 193—208.

Centre for Mathematics and Mechanics
1090 Sofia P. O. Box 373

Received 10. 5. 1982