ON THE APPROXIMATION OF DISTRIBUTIONS OF SUMS OF INDEPENDENT RANDOM VECTORS

A.Yu. Zaïtsev

Let $B_k$ be the σ-field of Borel subsets of the Euclidean space $\mathbb{R}^k$, $\mathcal{F}_k$ be the set of all probability measures on $B_k$, $\mathcal{D}_k$ be the set of infinitely divisible distributions in $\mathcal{F}_k$. The notation $x \in \mathbb{R}^k$ will indicate that $x = (x_1, \ldots, x_k)$ with $x_j \in \mathbb{R}$, $j = 1, \ldots, k$. For $x, y \in \mathbb{R}^k$ we write $x < y$ if $x_j < y_j$ for all $j = 1, \ldots, k$. For the $\varepsilon$-neighbourhood of a set $X \subseteq \mathbb{R}^k$ we use the notation $X^\varepsilon = \{y \in \mathbb{R}^k : \inf_{x \in X} ||x - y|| < \varepsilon\}$. We also denote: $1 = (1, 1, \ldots, 1) \in \mathbb{R}^k$, $\Phi \in \mathcal{F}_k$ the probability distribution concentrated at zero, $\Phi(F) \in \mathcal{F}_k$ the Gaussian distribution having the same mean and the same covariance operator as a given $F \in \mathcal{F}_k$, $\phi(F) = e^{-\frac{1}{2} \sum_{m=0}^\infty F^m/m!}$ (products and powers of measures are understood in the convolution sense), $\mathcal{F}(x) = \mathcal{F}\{\{u \in \mathbb{R}^k : u < x\}\}$. The symbols $C_1, C_2, \ldots$ will be used to denote positive constants depending only on the dimension $k$.

We shall estimate the following characteristics of proximity of distributions $F, G \in \mathcal{F}_k$: the uniform distance

$$\rho(F, G) = \sup_{x \in \mathbb{R}^k} |F(x) - G(x)|;$$

the multidimensional analogue of the Lévy distance

$$L(F, G) = \inf \{\varepsilon : F(x - \varepsilon 1) - \varepsilon \leq G(x) \leq F(x + \varepsilon 1) + \varepsilon \text{ for all } x \in \mathbb{R}^k\};$$

the Lévy–Prokhorov distance

$$\pi(F, G) = \inf \{\varepsilon : F\{X\} \leq G\{X^\varepsilon\} + \varepsilon, G\{X\} \leq F\{X^\varepsilon\} + \varepsilon \text{ for all } X \in B_k\}$$

and for $\lambda > 0$

$$L(F, G; \lambda) = \sup_{x \in \mathbb{R}^k} \max \{F(x + 0) - G(x + \lambda 1), G(x + 0) - F(x + \lambda 1)\};$$

$$\pi(F, G; \lambda) = \sup_{X \in B_k} \max \{F\{X\} - G\{X^\lambda\}, G\{X\} - F\{X^\lambda\}\}.$$
where \( U_i, V_i \in \mathcal{F}_k \),

\[
0 \leq p_i \leq p = \max_{1 \leq i \leq n} p_i \leq 1 ,
\]

the distributions \( U_i \) are such that

\[
U_i \{ \{ x \in \mathbb{R}^k : \| x \| \leq \tau \} \} = 1, \quad \int x U_i(\,dx\,) = 0
\]

and the distributions \( V_i \) are arbitrary. Similar representations of distributions arise quite often when we wish to perform a truncation of independent random summands. The problem is to estimate the closeness of the distribution \( F \) to the set \( D_k \) in terms of the parameters \( p \) and \( \tau \). The one-dimensional variant of this problem was considered in a slightly different form for the first time by Kolmogorov [5], see also [4], [6], [7] and [14]. Some infinite dimensional results were obtained by Bakšys and Paulauskas [3].

Let

\[
W_i = (1-p_i)U_i + p_i E, \quad G_i = (1-p_i)E + p_i V_i ,
\]

\[
W = \prod_{i=1}^n W_i, \quad G = \prod_{i=1}^n G_i, \quad W^* = \prod_{i=1}^n e(W_i), \quad G^* = \prod_{i=1}^n e(G_i).
\]

We shall use the following approximating distributions:

\[
D_1 = WG, \quad D_2 = \Phi(W)G, \quad D_3 = W^*G,
\]

\[
D_4 = WG^*, \quad D_5 = \Phi(W)G^*, \quad D_6 = W^*G^* = \prod_{i=1}^n e(F_i).
\]

Note that the distributions \( D_5 \) and \( D_6 \) are infinitely divisible. Moreover, \( D_6 \) is an accompanying infinitely divisible distribution for \( F = \prod_{i=1}^n F_i \).

**Theorem 1.** (see [10]). The following inequalities hold true:

\[
\pi(F, D_j) \leq C_1(p + \tau(\| \ln \tau \| + 1)), \quad j = 1, 2, 3;
\]

\[
\pi(F, D_j) \leq \sum_{i=1}^n p_i^2 + C_2(p + \tau(\| \ln \tau \| + 1)), \quad j = 4, 5, 6.
\]

If the measures \( V_i \) are identical for all \( i = 1, \ldots, n \) then

\[
\pi(F, D_j) \leq C_3(p + \tau(\| \ln \tau \| + 1)), \quad j = 4, 5, 6.
\]

The characteristics \( \tau(F, D_j, \lambda) \) may be estimated in a similar way replacing \( \tau(\| \ln \tau \| + 1) \) by \( \exp(-C_4 \lambda/\tau) \) on the right hand sides of the corresponding inequalities.
On the approximation of distributions

Theorem 2. For $L(\cdot, \cdot)$ the following inequalities are valid:

\[
L(F, D_j) \leq C_5 (p + \tau(|\ln \tau| + 1)), \quad j = 4, 5, 6,
\]

and for any $\lambda > 0$

\[
L(F, D_j; \lambda) \leq C_6 (p + \exp(-C_7 \lambda/\tau)), \quad j = 4, 5, 6.
\]

**Corollary.** Let $\varepsilon > 0$ and the distributions $F_i \in \mathcal{F}_k$ be such that $L(F_i, E) \leq \varepsilon$ for all $i = 1, \ldots, n$. Let $F = \prod_{i=1}^{n} F_i$. Then there exists a distribution $D \in \mathcal{D}_k$ such that

\[
L(F, D) \leq C_8 \varepsilon (|\ln \varepsilon| + 1).
\]

One-dimensional versions of the results formulated above and the history of the problem may be found in [2], [14]. A slightly weakened formulation of Theorem 1 was proved in [10]. Since $L(F, D_j) \leq \pi(F, D_j)$, the inequality (7) is also valid for $j = 1, 2, 3$. This follows from (4). Note that the term $\sum_{i=1}^{n} p_i^2$ cannot be removed from the right hand side of (5) without additional assumptions. This means that in general $L(\cdot, \cdot)$ cannot be replaced in (7) by $\pi(\cdot, \cdot)$.

We give below a plan of the proof of the inequality (7). In [10] it was proved that

\[
\pi(F, D_1) \leq C_8 (p + \tau(|\ln \tau| + 1)).
\]

Furthermore, from Theorems 1.1 and 1.2 [13] or from Theorem 1.1 [11] it follows that

\[
\pi(W, \Phi(W)) \leq C_9 \pi(|\ln \tau| + 1),
\]

\[
\pi(W^*, \Phi(W^*)) \leq C_{10} \pi(|\ln \tau| + 1).
\]

Taking into account the identity $\Phi(W) = \Phi(W^*)$ and well-known properties of the Lévy–Prokhorov distance, it can be easily derived from (11), (12) that

\[
\pi(D_1, D_2) \leq C_9 \pi(|\ln \tau| + 1),
\]

\[
\pi(D_4, D_5) \leq C_9 \pi(|\ln \tau| + 1),
\]

\[
\pi(D_2, D_3) \leq C_{10} \pi(|\ln \tau| + 1),
\]

\[
\pi(D_5, D_6) \leq C_{10} \pi(|\ln \tau| + 1).
\]

In order to estimate the closeness of $D_j$ and $D_{j+3}$, $j = 1, 2, 3$, we need the inequality

\[
\rho(G, G^*) \leq C_{11} p.
\]
A one-dimensional variant of this inequality was proved in [9] with the help of the so-called triangle function method introduced and for the first time applied by Arak [1], see also [2]. Note, however, that in order to prove the inequality (17) in multidimensional case it is necessary to revise the methods to be applied. The technical details connected with these changes in methods are discussed in [12].

It follows from (17) that

\[ \max_{j=1,2,3} \rho(D_j, D_{j+3}) \leq \rho(G, G^*) \leq C_{11} \rho. \]

Note that for any \( H_1, H_2 \in \mathcal{F}_k \) we have \( L(H_1, H_2) \leq \pi(H_1, H_2) \), \( L(H_1, H_2) \leq \rho(H_1, H_2) \). Therefore, the inequality (7) can be easily deduced from (10), (13)–(16), (18) with the help of the triangle inequality.

In spite of the fact that the inequality (8) seems to be essentially more general in comparison with (7), it can easily be derived from (7) by means of variation of a normalizing constant (see, for example, [2], [11], [13], [14]).

The method used in the proof of the inequality (17) gives a possibility to obtain other interesting results. For example, we formulate below two theorems about the closeness of multidimensional symmetric distributions. The set of symmetric distributions, i.e. distributions \( H \in \mathcal{F}_k^s \) for which \( H\{X\} = H\{-X\} \) for all \( X \in B_k \) will be denoted by \( \mathcal{F}_k^s \).

**Theorem 3.** Let \( h(t), t \in \mathbb{R}^k \), be the characteristic function of a distribution \( H \in \mathcal{F}_k^s \). Suppose that \( h(t) \) satisfies the inequality \( h(t) \geq -1 + \alpha, \alpha > 0, \) for all \( t \in \mathbb{R}^k \). Then for any natural number \( n \)

\[ \rho(H^n, H^{n+1}) \leq C_{12} (n^{-1} + \exp(-n\alpha + C_{13} \ln^3 n)). \]

**Theorem 4.** For any natural numbers \( m, n \) the following inequalities hold true:

\[ \sup_{H \in \mathcal{F}_k^s} \rho(H^n, (e(H))^n) \leq C_{14} n^{-1/2}, \]

\[ \sup_{H \in \mathcal{F}_k^s} \rho(H^n, H^{n+2}) \leq C_{15} n^{-1}, \]

\[ \sup_{H \in \mathcal{F}_k^s} \rho(H^n, H^{n+2m}) \leq C_{15} mn^{-1}, \]

\[ \sup_{H \in \mathcal{F}_k^s} \rho(H^n, H^{n+2m+1}) \leq C_{16} n^{-1/2} + C_{17} mn^{-1} \]

and, consequently,

\[ \sup_{m \leq \sqrt{n}} \sup_{H \in \mathcal{F}_k^s} \rho(H^n, H^{n+m}) \leq C_{18} n^{-1/2}. \]

**Remark.** For \( m = 0 \) the inequality (19) was proved in [8], and it is optimal with respect to the order. The corresponding counterexample is given by the
On the approximation of distributions

one-dimensional symmetric distribution $H$ concentrated on two points $\{-1, 1\}$: $H\{\{-1\}\} = H\{\{1\}\} = 1/2$. The one-dimensional variants of Theorems 3 and 4 are contained already in [2].

References


Leningrad Branch of Steklov Mathematical Institute
Fontanka 27
191011 Leningrad
USSR