Evgueni I. Gordienko

Comparing the distributions of sums of independent random vectors

Kybernetika, Vol. 41 (2005), No. 4, [519]--529

Persistent URL: http://dml.cz/dmlcz/135672

KYBERNETIKA — VOLUME 41 (2005), NUMBER 4, PAGES 519 – 529

COMPARING THE DISTRIBUTIONS OF SUMS OF INDEPENDENT RANDOM VECTORS

EVGUENI GORDIENKO

Let $(X_n, n \geq 1), (\tilde{X}_n, n \geq 1)$ be two sequences of i.i.d. random vectors with values in $\mathbb{R}^k$ and $S_n = X_1 + \cdots + X_n$, $\tilde{S}_n = \tilde{X}_1 + \cdots + \tilde{X}_n$, $n \geq 1$. Assuming that $EX_1 = E\tilde{X}_1$, $E|X_1|^2 < \infty$, $E|\tilde{X}_1|^{k+2} < \infty$ and the existence of a density of $\tilde{X}_1$ satisfying the certain conditions we prove the following inequalities:

$$v(S_n, \tilde{S}_n) \leq c \max\{v(X_1, \tilde{X}_1), \zeta_2(X_1, \tilde{X}_1)\}, \quad n = 1, 2, \ldots,$$

where v and ζ_2 are the total variation and Zolotarev's metrics, respectively.

Keywords: sum of random vectors, the total variation distance, bound of closeness, Zolotarev's metric, characteristic function

AMS Subject Classification: 60G50, 60F99

1. INTRODUCTION

We consider two sequences of independent and identically distributed random vectors $X_1, X_2, \ldots$ and $\tilde{X}_1, \tilde{X}_2, \ldots$ taking values from k-dimensional Euclidian space $\mathbb{R}^k$. Let

$$S_n = X_1 + \cdots + X_n, \; \tilde{S}_n = \tilde{X}_1 + \cdots + \tilde{X}_n, \quad n = 1, 2, \ldots,$$

and $X, \tilde{X}$ stand for generic random vectors distributed, correspondingly, as X_n and $\tilde{X}_n$.

We are concerned with upper bounds for the total variation distance $v(S_n, \tilde{S}_n)$, that (bounds) can be written as follows:

$$v(S_n, \tilde{S}_n) \leq c\mu(X, \tilde{X}), \quad n = 1, 2, \ldots, \tag{1.1}$$

where μ is a suitable probability metric.

The problem of estimation of $v(S_n, \tilde{S}_n)$ (or, of some other distance between the distributions of S_n and $\tilde{S}_n$) arises at least in two areas of probability theory. The first relates to estimating the rate of convergence in the central limit theorem, where both random vectors $\tilde{S}_n$ and $\tilde{X}$ are normally distributed, and the basic hypotheses are that

$$EX = E\tilde{X}, \quad M(X) = M(\tilde{X}), \quad E|X|^3 < \infty, \tag{1.2}$$

where $M(X)$ is the covariance matrix of X. In this set up inequalities (1.1) with $c = c'/\sqrt{n}$ are proved, provided that certain "smoothness" conditions are imposed on the distribution of X. To mention some examples of such results we refer the reader to [11–14].

Another field where the bounds as in (1.1) can be useful is the study of stability (continuity) of applied stochastic models involving some scheme of summation of random variables or random vectors. Examples of such models appear in analysis of important processes in queueing, insurance, finances, reliability, storage, etc. (see, e.g. [2, 6–9]). Often these models enclose, so-called, geometric sums with one-dimensional summands. Some results on stability (continuity) of distributions of such sums are given in [5, 6, 8]. In paper [10] the "inverse" stability problem is considered, when the deviations of the distributions of summands is estimated by the deviations of the distributions of a sum of random variables. Concerning sums of random vectors on $\mathbb{R}^k$ $(k > 1)$ the only example of continuity bounds we know is the estimation of Zolotarev's distance between sums of random vectors with a geometrically distributed number of summands and an appropriately defined k-dimensional exponential random vector (see [8]).

In the stability problem setting the distribution of the random vector $\tilde{X}$ is interpreted as a known (available) approximation to an unknown distribution of the random vector X. For this reason it is desirable to have the constant c in (1.1) independent of the distribution of X.

In contrast to (1.2) we will suppose the equality of the first moments: $EX = E\tilde{X}$ admitting distinct covariance matrices of X and of $\tilde{X}$. Assuming also the boundedness of some power moments and the existence of an "enough smooth" density of the known random vector $\tilde{X}$ (see Assumptions 2.1 and 2.2) we prove that

$$v(S_n, \tilde{S}_n) \le c\, \mu(X, \tilde{X}), \quad n = 1, 2, \dots, \tag{1.3}$$

provided that

$$\mu(X, \tilde{X}) \le (2c)^{-1}, \tag{1.4}$$

where $\mu = \max(v, \zeta_2)$ and ζ_2 is Zolotarev's metric of order 2 (see the definition in Section 2). The constant c in (1.3), (1.4) does not depend on n and it is completely determined by characteristics of the density of $\tilde{X}$.

We consider a simple example of application of bound (1.3) to estimate the stability of the joint distribution of a total asset return and a total trading volume during a "stable interval" of the functioning of a stock market. The corresponding model was discussed in [8].

2. NOTATIONS, DEFINITIONS AND ASSUMPTIONS

For $x = (x_1, \dots, x_k) \in \mathbb{R}^k$ let

$$|x| := \left(\sum_{i=1}^{k} x_i^2 \right)^{1/2}, \qquad \|x\| := \sum_{i=1}^{k} |x_i|.$$

Let $g : \mathbb{R}^k \to \mathbb{R}$ be a function having all partial derivatives of order 2. We will write:

$$Dg := (D_1 g, \dots, D_k g), \quad D^2 g := \{D_{ij} g : i, j = 1, \dots, k\},$$

where

$$D_i g(x) = \frac{\partial}{\partial x_i} g(x), \quad D_{ij} g(x) = \frac{\partial^2}{\partial x_i \partial x_j} g(x).$$

We will use also the following notation:

$$Dg \in \mathbb{L}_1 \quad \text{if} \quad D_i g \in \mathbb{L}_1(\mathbb{R}^k), \quad i = 1, \dots, k;$$
$$D^2 g \in \mathbb{L}_1 \quad \text{if} \quad D_{ij} g \in \mathbb{L}_1(\mathbb{R}^k), \quad i, j = 1, \dots, k;$$
$$\|D^2 g\|_{\mathbb{L}_1} := \max_{1 \le i,j \le k} \|D_{ij} g\|_{\mathbb{L}_1(\mathbb{R}^k)}. \tag{2.1}$$

In what follows we will need three probability metrics $\boldsymbol{v}, \zeta_2$ and μ measuring the distances between distributions of random vectors.

— The total variation metric is defined by the formula:

$$\boldsymbol{v}(X, Y) := 2 \sup\{|P(X \in B) - P(Y \in B)| : B \subset \mathbb{R}^k \text{ is a Borel set}\}$$

— Zolotarev's metric of order 2 is defined as follows (see, e. g. [8, 14]).

$$\zeta_2(X, Y) := \sup\{|E\varphi(X) - E\varphi(Y)| : \varphi \in \mathcal{D}_2\}, \tag{2.2}$$

where

$$\mathcal{D}_2 := \{\varphi : \mathbb{R}^k \to \mathbb{R} : \|D\varphi(x) - D\varphi(y)\| \le \|x - y\|, x, y \in \mathbb{R}^k\}.$$

It is well known (see [8, 14]) that $\zeta_2(X, Y) < \infty$ if $EX = EY$ and $E|X|^2$, $E|Y|^2 < \infty$.

— Metric μ is defined as the following combination of $\boldsymbol{v}$ and ζ_2:

$$\mu(X, Y) := \max\{\boldsymbol{v}(X, Y), \zeta_2(X, Y)\}.$$

Now using the notation of the beginning of Introduction we set hypotheses which we need to prove the main inequalities (3.4) given in Section 3. We emphasize that apart from condition (2.3) the restrictions we use are concerned with the distribution of a random vector $\tilde{X}$ which is supposed to be "known" in the problem of evaluation of stability.

Assumption 2.1.

(a) $EX = E\tilde{X}, \quad E|X|^2 < \infty.$ (2.3)

(b) $E|\tilde{X}|^{k+2} < \infty$ and the covariance matrix of $\tilde{X}$ denoted throughout by M is positive definite.

Assumption 2.2. The random vector $\tilde{X}$ has a density f with respect to the Lebesgue measure on $\mathbb{R}^k$. Moreover, there exists an integer $s \geq 1$ such that the density f_s of $\tilde{S}_s = \tilde{X}_1 + \cdots + \tilde{X}_s$ satisfies the following conditions:

(a) f_s is bounded and has all second partial derivatives;

(b) Df_s is bounded and belongs to $\mathbb{L}_1$;

(c) $D^2 f_s$ is bounded, continuous, it belongs to $\mathbb{L}_1$, and moreover there exists $\alpha > 0$ such that

$$\int_{|x|>\alpha n} |D_{ij} f_s(x)| \, dx = O\left(\frac{1}{n}\right) \quad \text{as } n \to \infty \tag{2.4}$$

for all $1 \leq i,\ j \leq k$.

3. MAIN RESULT AND EXAMPLE

Let us denote by g_n the density of the random vector $\frac{1}{\sqrt{n}}(\tilde{X}_1 + \cdots + \tilde{X}_n)$, $n \geq 1$.

It is proved in Lemma 4.1 that under Assumption 2.1 and 2.2

$$d := \sup_{n \geq s} \|D^2 g_n\|_{\mathbb{L}_1} < \infty. \tag{3.1}$$

Set

$$c = \max\{2s - 1, 5dk\}, \tag{3.2}$$

where s is the integer from Assumption 2.2.

Theorem 3.1. Suppose that Assumptions 2.1 and 2.2 hold. If

$$\mu(X, \tilde{X}) \leq (2c)^{-1} \tag{3.3}$$

then

$$v(S_n, \tilde{S}_n) \leq c\mu(X, \tilde{X}), \quad n = 1, 2, \ldots . \tag{3.4}$$

From (3.1) and (3.2) it follows that the definition of the constant c does not involve anything related to the distribution of X. On the other hand, c is specified by the distribution of $\tilde{X}$ in a rather intricate way (see (3.1)). Unfortunately, the proof of Lemma 4.1 does not provide with an affective method for evaluating d in (3.1). Nevertheless, in some particular cases the densities g_n and their derivatives $D^2 g_n$ can be calculated explicitly. This allows to estimate the constant d by computer calculations. Indeed, one can deduce from the proof of Lemma 4.1 that

$$\|D^2 g_n\|_{\mathbb{L}_1} \to \|D^2 \varphi\|_{\mathbb{L}_1}$$

as $n \to \infty$, where φ is the normal density with zero mean and the covariance matrix M. Thus, to bound d it suffices to choose "experimentally" an appropriate integer $n_* > s$ and to calculate numerically the values of $\|D^2 g_n\|_{\mathbb{L}_1}$ for $n = s, s+1, \ldots, n_*$. Numerical experiments with gamma-densities and uniform densities (for $k = 1$ and k=2) displayed that for these densities the sequence $\{\|D^2 g_n\|_{\mathbb{L}_1}, n \geq n'\}$ is decreasing and $n' - s < 10$. Thus finding the above n_* is plain.

For instance, let us consider a two-dimensional random vector $\tilde{X}$ with independent components which have the same gamma-density $\frac{\lambda}{6}x^3 e^{-\lambda x}$. Then Assumption 2.2 is satisfied for $s = 1$. We have computed that $d < 0.45$ for $\lambda = 1$ and by simple calculations we get that in (3.3), (3.4) $c < 4.5\lambda^2$.

To illustrate an application of bound (3.4) we estimate the stability of one model studied in [8]. In this model the functioning of stock market is considered in discrete time (say, days) $n = 1, 2, \ldots$ during an interval of a random length ν. This interval corresponds to a period of "relatively stable prices" (till some dramatically change of prices). Let us denote by ξ_n the asset return at the *nth* day and by η_n the trading volume at the same day. In analysis of stock market a matter of interest is the distribution of the random vector $Y = (\sum_{n=1}^{\nu} \xi_n,\ \sum_{n=1}^{\nu} \eta_n)$ which represents the total return and the total trading volume during the stable interval. It is supposed that $(\xi_1, \eta_1), (\xi_2, \eta_2), \ldots$ are i.i.d. random vectors. Let us imagine that the distribution of a random vector $X = (\xi_1, \eta_1)$ is uncertain and the distribution of some random vector $\tilde{X} = (\tilde{\xi}_1, \tilde{\eta}_1)$ is used as an available approximation. Let $\tilde{X}$ satisfy Assumptions 2.1, 2.2 and the random variable ν be independent of $\{(\xi_n, \eta_n),\ n \geq 1\}$. Using the random vector $\tilde{Y} := \left(\sum_{n=1}^{\nu} \tilde{\xi}_n, \sum_{n=1}^{\nu} \tilde{\eta}_n\right)$ to approximate Y we get by the total probability formula:

$$v(Y, \tilde{Y}) \leq \sum_{n=1}^{\infty} v \left[\left(\sum_1^n \xi_k,\ \sum_1^n \eta_k \right), \left(\sum_1^n \tilde{\xi}_k, \sum_1^n \tilde{\eta}_k \right) \right] P(\nu = n).$$

Finally, we obtain from (3.4) that

$$v(Y, \tilde{Y}) \leq c\mu(X, \tilde{X}),$$

if condition (3.3) holds.

4. THE PROOF OF THEOREM 3.1

Let f_n and g_n be the densities of random vectors $\tilde{S}_n = \tilde{X}_1 + \cdots + \tilde{X}_n$ and $Z_n := \frac{1}{\sqrt{n}} \tilde{S}_n$, respectively, and let $\varphi_n = \varphi_n(t)$, $t \in \mathbb{R}^k$ denote the characteristic function of Z_n $(n \geq 1)$. Particularly, $\varphi_1 =: \varphi$ is the characteristic function of $\tilde{X}$.

Since $v(X + a,\ Y + a) = v(X, Y)$ for any random vectors X, Y and $a \in \mathbb{R}^k$ we suppose without loss of generality that $EX = E\tilde{X} = 0$ (see Assumption 2.1, (a)). The next lemma is the key step in proving Theorem 3.1. Throughout the proof of this lemma we use the letter c to denote finite constants, possibly distinct in different places.

Lemma 4.1. Under Assumptions 2.1 and 2.2

$$d := \sup_{n \geq s} \|D^2 g_n\|_{\mathbb{L}_1} < \infty.$$

Proof. In the first place we show the existence of an integer $\ell \geq s$ such that $|t|^2\varphi_\ell(t) \in \mathbb{L}_1(\mathbb{R}^k)$ and, consequently,

$$|t|^2\varphi_n(t), \quad \varphi_n \in \mathbb{L}_1(\mathbb{R}^k) \quad \text{for} \quad n \geq \ell. \tag{4.1}$$

For arbitrary, but fixed i ($1 \leq i \leq k$) and $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \in \mathbb{R}^{k-1}$ let $v_i(x_i) := \frac{\partial}{\partial x_i} f_s(x_1, \ldots, x_n)$. It is easy to see from Assumption 2.2, (b) that $v_i \in \mathbb{L}_1(\mathbb{R})$. Therefore we can integrate by parts with respect to x_i in the definition of a characteristic function:

$$\begin{aligned}\varphi_s(\sqrt{s}t) &= \int_{\mathbb{R}^k} f_s(x)e^{i\langle t,x\rangle}\,\mathrm{d}x \\ &= \int_{\mathbb{R}^{k-1}} e^{i\langle t',x'\rangle}\,\mathrm{d}x' \int_{\mathbb{R}} f_s(\ldots, x_i, \ldots)e^{it_ix_i}\,\mathrm{d}x_i \\ &= \int_{\mathbb{R}^{k-1}} e^{i\langle t',x'\rangle}\,\mathrm{d}x' \frac{1}{it_i} \int_{\mathbb{R}} v_i(x_i)e^{it_ix_i}\,\mathrm{d}x_i.\end{aligned} \tag{4.2}$$

Again, appealing to Assumption 2.2, (b) we deduce from (4.2) that $|\varphi_s(\sqrt{s}t)| \leq \frac{c_i}{|t_i|}$. Consequently, $|\varphi_s(\sqrt{s}t)| \leq \frac{c}{|t|}$, $t \in \mathbb{R}^k$.

Thus $|t|^2\varphi_s^m \in \mathbb{L}(\mathbb{R}^k)$ for an enough large m and we can choose $\ell = ms$ to get (4.1).

Condition (4.1) allows to write down the inverse Fourier transform:

$$g_n(x) = \frac{1}{(2\pi)^k} \int_{\mathbb{R}^k} \varphi_n(t)e^{-i\langle t,x\rangle}\,\mathrm{d}t$$

and to differentiate under an integral sign in this equality (see Appendix A, Corollary A.15 in [4]). Consequently, one can easily show that for every $n \geq \ell$, $1 \leq i,\ j \leq k$ the function $-t_it_j\varphi_n(t)$ is a Fourier transform of $D_{ij}g_n$.

As a next step we show how to approximate $D_{ij}g_n$ by a polynomial times a normal density. In the rest of the proof of the lemma we fix an arbitrary pair of integers i, j ($1 \leq i, j \leq k$). We are going to prove the existence of polynomials $\tilde{P}_0, \tilde{P}_1, \ldots, \tilde{P}_k$ and a constant c such that

$$\sup_{x\in\mathbb{R}^k} \left| D_{ij}g_n(x) - \sum_{m=0}^{k} n^{-m/2}\tilde{P}_m(x)\varphi_{0,M}(x) \right| \leq \frac{c}{n^{k/2}}, \quad n = \ell, \ell+1, \ldots \tag{4.3}$$

where $\varphi_{0,M}$ is the density of the normal distribution with zero mean and the covariance matrix M. (This matrix appeared in Assumption 2.1, (b).)

To verify (4.3) we observe that Assumption 2.1 (b), together with the condition $\varphi_n \in \mathbb{L}_1(\mathbb{R}^k)$, $n \geq \ell$ guarantees the hypotheses of Theorem 19.2 Ch.4 in [3]. In particular this theorem yields the following asymptotic expansion of the density g_n:

$$g_n(x) = \sum_{m=0}^{k} n^{-m/2}P_m(x)\varphi_{0,M}(x) + o\left(\frac{1}{n^{k/2}}\right), \quad x \in \mathbb{R}^k$$

(as $n \to \infty$), where $P_o, \dots, P_k$ are certain polynomials. Let us set

$$h_n(x) := g_n(x) - \sum_{m=0}^{k} n^{-m/2} P_m(x)\varphi_{0,M}(x), \quad x \in \mathbb{R}^k$$

and differentiate this equality to get the following:

$$D_{ij}h_n(x) = D_{ij}g_n(x) - D_{ij}\left[\sum_{m=0}^{k} n^{-m/2} P_m(x)\varphi_{0,M}(x)\right]. \tag{4.4}$$

Let q_n be the Fourier transform of $D_{ij}h_n$. In view of Lemma 7.2, Ch. 2 in [3] the Fourier transform of $\sum_{m=0}^{k} n^{-m/2} P_m \varphi_{0,M}$ is a function of the form:

$$\sum_{m=0}^{k} n^{-m/2} \hat{P}_m(t) \exp\left\{-\frac{1}{2}\langle t, Mt\rangle\right\}, \quad t \in \mathbb{R}^k,$$

where $\hat{P}_0, \dots, \hat{P}_k$ are certain polynomials. Therefore $q_n(t) = -t_i t_j \hat{h}_n(t)$, where

$$\hat{h}_n(t) = \varphi_n(t) - \sum_{m=0}^{k} n^{-m/2} \hat{P}_m(t) \exp\left\{-\frac{1}{2}\langle t, Mt\rangle\right\}. \tag{4.5}$$

Since $q_n \in \mathbb{L}_1(\mathbb{R}^k)$ (see (4.1)) we get

$$D_{ij}h_n(x) = -\frac{1}{(2\pi)^k} \int_{\mathbb{R}^k} t_i t_j \hat{h}_n(t) e^{-i\langle t, x\rangle}\, dt, \quad x \in \mathbb{R}^k. \tag{4.6}$$

It was shown in the proof of Theorem 19.2, Ch.4 in [3] that there is a positive constant γ such that

$$\left|\hat{h}_n(t)\right| \le \frac{c}{n^{k/2}} \langle t, Mt\rangle^{3k} \exp\left\{-\frac{1}{4}\langle t, Mt\rangle\right\}$$

for all $|t| \le \gamma\sqrt{n}$. Thus

$$\int_{|t| \le \gamma\sqrt{n}} |t_i t_j|\, |\hat{h}_n(t)|\, dt \le \frac{c}{n^{k/2}}. \tag{4.7}$$

In view of (4.4), (4.6) and (4.7) to prove (4.3) it is enough to show that

$$\int_{|t| > \gamma\sqrt{n}} |t_i t_j|\, |\hat{h}_n(t)|\, dt = O\left(\frac{1}{n^{k/2}}\right)$$

or that $\int_{|t|>\gamma\sqrt{n}} |t_i t_j|\, |\varphi_n(t)|\, dt = O\left(\frac{1}{n^{k/2}}\right)$ as $n \to \infty$ (see 4.5)).

Since $\tilde{X}$ has a density we get that $\sup_{|t|>\gamma\sqrt{n}} \left|\varphi\left(\frac{t}{\sqrt{n}}\right)\right| \le \beta < 1$. Hence (see(4.1))

$$\begin{aligned}\int_{|t|>\gamma\sqrt{n}} |t_i t_j|\, |\varphi_n(t)|\, dt &= \int_{|t|>\gamma\sqrt{n}} |t_i t_j| \left|\varphi^{\ell}\left(\frac{t}{\sqrt{n}}\right)\right| \left|\varphi^{n-\ell}\left(\frac{t}{\sqrt{n}}\right)\right| dt \\ &\le \beta^{n-\ell} n^{\frac{k+2}{2}} \int_{\mathbb{R}^k} |\tau_i \tau_j|\, |\varphi^{\ell}(\tau)|\, d\tau \le \frac{c}{n^{k/2}}\end{aligned}$$

for a suitable constant c.

Now we are in position to prove that

$$\sup_{n\geq s} \|D_{ij}g_n\|_{\mathbb{L}_1(\mathbb{R}^k)} < \infty.$$

It is clear that $g_n(x) = n^{\frac{k}{2}} f_n(\sqrt{n}x)$ and

$$f_n(x) = \int_{\mathbb{R}^k} f_s(x-t) f_{n-s}(t)\,\mathrm{d}t \quad \text{for } n > s, \tag{4.8}$$

where f_n is the density of $\tilde{S}_n$. By virtue of Assumption 2.2 and Corollary A.14, Appendix A in [4] we can differentiate under an integral sign in (4.8). Hence (see Assumption 2.2, (c))

$$D_{ij} f_n(x) = \int_{\mathbb{R}^k} D_{ij} f_s(x-t) f_{n-s}(t)\,\mathrm{d}t \in \mathbb{L}_1(\mathbb{R}^k) \tag{4.9}$$

for all $n > s$ by Fubini's Theorem; also $D_{ij} f_s \in \mathbb{L}_1(\mathbb{R}^k)$. We have:

$$\|D_{ij}g_n\|_{\mathbb{L}_1(\mathbb{R}^k)} = \int_{|x|\leq 2\sqrt{n}\,\alpha} |D_{ij}g_n|\,\mathrm{d}x + \int_{|x|>2\sqrt{n}\,\alpha} |D_{ij}g_n|\,\mathrm{d}x, \tag{4.10}$$

where α is the constant from Assumption 2.2, (c). On the strength of (4.3) the first terms on the right-hand side of (4.10) are uniformly bounded in $n \geq \ell$. Let us show uniform boundedness of the second terms in (4.10). By (4.9) we have that

$$\begin{aligned} \int_{|x|>2\sqrt{n}\alpha} |D_{ij}g_n|\,\mathrm{d}x &= n \textstyle\int_{|z|>2\alpha n} |D_{ij}f_n(z)|\,\mathrm{d}z \\ &\leq n \int_{|z|>2\alpha n} \mathrm{d}z \int_{\mathbb{R}^k} |D_{ij} f_s(z-t)| f_{n-s}(t)\,\mathrm{d}t =: nI_n. \end{aligned} \tag{4.11}$$

Also

$$\begin{aligned} I_n &= \int_{|z|>2\alpha n} \mathrm{d}z \int_{|t|\leq\alpha n} |D_{ij} f_s(z-t)| f_{n-s}(t)\,\mathrm{d}t \\ &\quad + \int_{|z|>2\alpha n} \mathrm{d}z \int_{|t|>\alpha n} |D_{ij} f_s(z-t)| f_{n-s}(t)\,\mathrm{d}t =: \mathcal{J}_n + \mathcal{J}'_n. \end{aligned}$$

The inequalities $|z| > 2\alpha n$ and $|t| \leq \alpha n$ yield $|z-t| > \alpha n$. Hence

$$\mathcal{J}_n = \int_{|t|\leq\alpha n} f_{n-s}(t)\,\mathrm{d}t \int_{|z|>2\alpha n} |D_{ij} f_s(z-t)|\,\mathrm{d}z \leq \frac{c}{n} \tag{4.12}$$

due to (2.4).

On the other hand, since $D_{ij} f_s \in \mathbb{L}_1(\mathbb{R}^k)$ we get that

$$\begin{aligned} \mathcal{J}'_n \leq c \int_{|t|>\alpha n} f_{n-s}(t)\,\mathrm{d}t &= cP\Big(\Big(\textstyle\sum_1^{n-s} |\tilde{X}_i| > \alpha n\Big) \\ &\leq c\frac{E\left|\sum_1^{n-s} \tilde{X}_i\right|^{k+2}}{\alpha^{k+2} n^{k+2}} \leq c' \frac{(n-s)^{\frac{k+2}{2}}}{n^{k+2}} \leq \frac{c''}{n} \end{aligned} \tag{4.13}$$

on the strength of Assumption 2.1, (b) and the Rosenthal inequality (see [1]):

$$E\left|\sum_1^n \tilde{X}_i\right|^p \le c(k,p)\max\left\{\sum_1^n E|\tilde{X}_i|^p, \left(\sum_1^n E|\tilde{X}_i|^2\right)^{p/2}\right\}, \quad p>2.$$

Aggregating relations (4.10) – (4.13) completes the proof. □

Apart from the above lemma the proof of Theorem 3.1 is similar to the prove of Lemma 1 in [6] where distributions of sums of random variables were compared (i. e. $\mathbb{R}^k = \mathbb{R}$). This approach takes advantage of the "metric" techniques which is well-known in estimating the rate of convergence in the central limit theorem (see, e. g. [11, 12, 14]). That is why we would rather outline the main points of the proof than give a complete demonstration.

By the regularity of the metric v we get that $v(S_n, \tilde{S}_n) \le cv(X, \tilde{X})$, $n = 1, 2, \dots, s-1$, provided that $c \ge 2s-1$.

For $n \ge 2s$ let $m := [n/2] \ge s$ ($[\cdot]$ is the integer part). Using the triangle inequality and the following well-known inequality (see [11]):

$$v(X+U, Z+U) \le v(X,Z)v(U,V) + v(X+V, Z+V)$$

valued for any independent random vectors, X, Z, U, v, and choosing

$$X = X_1 + \dots + X_m, \quad U = X_{m+1} + \dots + X_n,$$
$$Z = \tilde{X}_1 + \dots + \tilde{X}_m, \quad V = \tilde{X}_{m+1} + \dots + \tilde{X}_n,$$

we deduce that

$$\begin{aligned} v(S_n, \tilde{S}_n) \le\ & v(X_1 + \dots + X_m, \tilde{X}_1 + \dots + \tilde{X}_m)\times \\ & \times v(X_{m+1} + \dots + X_n, \tilde{X}_{m+1} + \dots + \tilde{X}_n) + T_n' + T_n'', \end{aligned}$$

where

$$T_n' = v\left(\frac{\tilde{X}_1 + \dots + \tilde{X}_m}{\sqrt{n-m}} + \frac{\tilde{X}_{m+1} + \dots + \tilde{X}_n}{\sqrt{n-m}}, \frac{X_1 + \dots + X_m}{\sqrt{n-m}} + \frac{\tilde{X}_{m+1} + \dots + \tilde{X}_n}{\sqrt{n-m}}\right) \tag{4.15}$$

$$T_n'' = v\left(\frac{\tilde{X}_{m+1} + \dots + \tilde{X}_n}{\sqrt{m}} + \frac{\tilde{X}_1 + \dots + \tilde{X}_m}{\sqrt{m}}, \frac{X_{m+1} + \dots + X_n}{\sqrt{m}} + \frac{\tilde{X}_1 + \dots + \tilde{X}_m}{\sqrt{m}}\right). \tag{4.16}$$

To obtain the above form of T_n' and T_n'' we also used the fact that $v(aX, aY) = v(X,Y)$, $a \ne 0$.

To bound T_n' and T_n'' we make use of the following assertion.

Assume that X, Y and ξ are independent random vectors taking values in $\mathbb{R}^k$ and that ξ has a density f_ξ such that $D^2 f_\xi$ is bounded, continuous and belongs to $\mathbb{L}_1$. Then

$$v(X+\xi, Y+\xi) \leq k\|D^2 f_\xi\|_{\mathbb{L}_1} \zeta_2(X,Y), \tag{4.17}$$

where the norm $\|\cdot\|_{\mathbb{L}_1}$ was defined in (2.1) and ζ_2 is Zolotarev's metric of order 2 defined in (2.2).

A version of (4.17) was proved in [11] for ξ being a normally distributed random vector. In our more general case the proof of the above inequality can be carried out along the same lines. In view of Assumption 2.2, (c) and Lemma 4.1 we can apply inequality (4.17) to bound the right-hand sides of (4.15) and (4.16). Moreover, we take advantage of the so-called ideality properties of the metric ζ_2 (see [8, 12, 14]), i. e.

$$\zeta_2\left(a\sum_{i=1}^n X_i, a\sum_{i=1}^n Y_i\right) \leq a^2 \sum_{i=1}^n \zeta_2(X_i, Y_i), \quad a>0 \tag{4.18}$$

for independent random vectors $X_1, \ldots, X_n;\ Y_1, \ldots, Y_n$ with $EX_i = EY_i$ $(i = 1, 2, \ldots, n)$.

Combining (4.14) – (4.17) with the result of Lemma 4.1 we get the inequality:

$$\begin{aligned} v(S_n, \tilde{S}_n) &\leq v(S_m, \tilde{S}_m) v(S_{n-m}, \tilde{S}_{n-m}) \\ &+ kd\frac{[n/2]}{n-[n/2]}\zeta_2(X, \tilde{X}) + kd\frac{n-[n/2]}{[n/2]}\zeta_2(X, \tilde{X}), \end{aligned}$$

(see (3.1) for the definition of d).

Consequently,

$$v(S_n, \tilde{S}_n) \leq \mu(S_m, \tilde{S}_m)\mu(S_{n-m}, \tilde{S}_{n-m}) + 2.5kd\mu(X, \tilde{X}). \tag{4.19}$$

Making the induction assumption that

$$v(S_r, \tilde{S}_r) \leq c\mu(X, \tilde{X}), \qquad r \leq n-1 \tag{4.20}$$

we see that (4.19) would yield (4.20) with $r = n$ if we choose $c = \max\{2s-1, 5kd\}$ and require that $\mu(X, \tilde{X}) \leq (2c)^{-1}$. □

ACKNOWLEDGEMENT

The author would like to thank the referees for valuable suggestions.

(Received July 18, 2003.)

REFERENCES

[1] A. Araujo and E. Giné: The Central Limit Theorem for Real and Banach Valued Random Variables. Wiley, New York 1980.

[2] S. Asmussen: Applied Probability and Queues. Wiley, Chichester 1987.

[3] R. N. Bhattacharya and R. Ranga Rao: Normal Approximation and Asymptotic Expansions. Wiley, New York 1976.

[4] R. M. Dudley: Uniform Central Limit Theorems. Cambridge University Press, Cambridge 1999.
[5] E. I. Gordienko: Estimates of stability of geometric convolutions. Appl. Math. Lett. *12* (1999), 103–106.
[6] E. I. Gordienko and J. Ruiz de Chávez: New estimates of continuity in $M|GI|1|\infty$ queues. Queueing Systems Theory Appl. *29* (1998), 175–188.
[7] J. Grandell: Aspects of Risk Theory. Springer–Verlag, Heidelberg 1991.
[8] V. Kalashnikov: Geometric Sums: Bounds for Rare Events with Applications. Kluwer Academic Publishers, Dordrecht 1997.
[9] V. Kalashnikov and D. Konstantinidis: The ruin probability. Fund. Appl. Math. *2* (1996), 1055–1100 (in Russian).
[10] A. V. Prokhorov and N. G. Ushakov: On the problem of reconstructing a summands distribution by the distribution of their sum. Theory Probab. Appl. *46* (2002), 420–430.
[11] V. V. Senatov: Uniform estimates of the rate of convergence in the multi-dimensional central limit theorem. Theory Probab. Appl. *25* (1980), 745–759.
[12] V. V. Senatov: Qualitative effects in estimates for the rate of convergence in the central limit theorem in multidimensional spaces. Proc. Steklov Inst. Math. *215* (1996), 4, 1–237.
[13] Yu. V. Zhukov: On the accuracy of normal approximation for the densities of sums of independent identically distributed random variables. Theory Probab. Appl. *44* (2000), 785–793.
[14] V. Zolotarev: Ideal metrics in the problems of probability theory. Austral. J. Statist. *21* (1979), 193–208.

Evgueni Gordienko, Universidad Autónoma Metropolitana – Unidad Iztapalapa, Av. San Rafael Atlixco # 186, Colonia Vicentina, México 09340, D.F. Mexico.
e-mail: gord@@xanum.uam.mx