The Stein-Tikhomirov Method and Berry-Esseen Inequality for Sampling Sums from a Finite Population of Independent Random Variables

Shakir K. Formanov and Tamara A. Formanova

Abstract We present a simplified version of the Stein-Tikhomirov method realized by defining a certain operator in class of twice differentiable characteristic functions. Using this method, we establish a criterion for the validity of a nonclassical central limit theorem in terms of characteristic functions, in obtaining of classical Berry-Esseen inequality for sampling sums from finite population of independent random variables.

Keywords Stein-Tikhomirov method • Distribution function • Characteristic function • Independent random variables • Berry-Esseen inequality • Sampling sums from finite population

Mathematics Subject Classification (2010): 60F05

1 The Stein-Tikhomirov Method and Nonclassical CLT

Suppose that $F(x)$ is an arbitrary distribution function and

$$\Phi(x) = \frac{1}{2\pi} \int_{-\infty}^{x} e^{-u^2/2} du$$

S.K. Formanov (✉)
National University of Uzbekistan, Tashkent, Uzbekistan
e-mail: shakirformanov@yandex.ru

T.A. Formanova
Tashkent Institute of Motor Car and Road Engineers, Tashkent, Uzbekistan
e-mail: fortamara@yandex/ru

A.N. Shiryaev et al. (eds.), *Prokhorov and Contemporary Probability Theory*, Springer Proceedings in Mathematics & Statistics 33,
DOI 10.1007/978-3-642-33549-5_14,

is the standard distribution function for the normal law. In [9] Stein proposed a universal method for estimating the quantity

$$\delta = \sup_x |F(x) - \Phi(x)|,$$

based on the following arguments. Suppose that $h(u)$ is a bounded measurable function on the line and

$$\Phi h = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(u) e^{-u^2/2} du.$$

Consider the function $g(\cdot)$ which is a solution of the differential equation

$$g'(u) - ug(u) = h(u) - \Phi h. \tag{1}$$

Suppose that ζ is a random variable with distribution function

$$P(\zeta < x) = F(x).$$

Setting

$$h(u) = h_x(u) = I_{(-\infty,x)}(u)$$

in (1), we have

$$F(x) - \Phi(x) = E\left[g'(\zeta) - \zeta g(\zeta)\right]. \tag{2}$$

Thus, the problem of estimating δ can be reduced to that of estimating the difference of the expectations

$$\left|Eg'(\zeta) - E\zeta g(\zeta)\right|.$$

Also note that for the case in which the random variable ζ has normal distribution, the right-hand side of (2) vanishes. Using this method, Stein [9] obtained an estimate of the rate of convergence in the central limit theorem for stationary (in the narrow sense) sequences of random variables satisfying the strong mixing conditions (in the sense of Rosenblatt). Moreover, for the summands eighth-order moments must exits. In his paper, Stein stated his belief that his method is hardly related to that of characteristic functions.

In [10, 11] Tikhomirov refuted Stein's suggestion. He showed that a combination of Stein's ideas with the method of characteristic functions allows one to obtain the best possible estimates of the rate of convergence in the central limit theorem for sequences of weakly dependent random variables for less stringent conditions on the moments. He also used to best advantage the ideas [9] underling the proposed new method. The combination of methods our lined in [9, 10], later became known as the Stein-Tikhomirov method.

In the present paper, it will be shown that the arguments used in applying the Stein-Tikhomirov method can be considerably simplified. Thus will be demonstrated in the course of the proof of a nonclassical central limit theorem. Which can be called the generalized Lindeberg-Feller theorem.

Suppose that

$$X_{n1}, X_{n2}, \cdots$$

is a sequence of independent random variables constituting the scheme of a series of experiments and

$$S_n = X_{n1} + X_{n2} + \cdots, n = 1, 2, \cdots$$

with a possibly infinite number of terms in each sum. Set

$$EX_{nj} = 0, \quad EX_{nj}^2 = \sigma_{nj}^2, \quad j = 1, 2, \cdots$$

and

$$\sum_j \sigma_{nj}^2 = 1. \tag{3}$$

In what follows, condition (3) is assumed to be satisfied. As is well known, in the theory of summation of independent random variables an essential role is played by the condition of uniform infinite smallness of the summands

$$\lim_{n\to\infty} \sup_j P\left(\left|X_{nj}\right| \geq \varepsilon\right) = 0 \tag{4}$$

for any $\varepsilon > 0$.

The constraint (4) is needed if we want to make the limiting law for the distribution of the sum S_n insensitive to the behavior of individual summands. But in finding conditions for the conditions for the convergence of the sequence of distributions functions

$$F_n(x) = P(S_n < x)$$

for any given law it is not necessary to introduce constraints of type (4). Following Zolotarev, limit theorems making no use of condition (4) are said to be *nonclassical*. As was noted in the monograph "theory of summation of independent random variables", the ideas underlying the nonclassical approach go back to P.Lévy, who studied various versions of the central limit theorem.

In [7], Rotar' proved the following theorem, which is generalization of the classical Lindeberg-Feller theorem.

Theorem A. *In order that*

$$\sup_x |F_n(x) - \Phi(x)| \to 0$$

as $n \to \infty$, it is necessary and sufficient that for any $\varepsilon > 0$ the following relation hold:

$$R_n(\varepsilon) = \sum_i \int_{|x|>\varepsilon} |x| \left| F_{nj}(x) - \Phi_{nj}(x) \right| \to 0, \tag{5}$$

where

$$F_{nj}(x) = P\left(X_{nj} < x\right), \quad \Phi_{nj}(x) = \Phi\left(\frac{x}{\sigma_{nj}}\right).$$

Note that this version of Theorem A is not given in [7], but it can be obtained by combining Propositions 1 and 2 from [7].

The numerical characteristic $R_n(\varepsilon)$ defined in (5) is universal; it and its analogs have been used for some time in the "nonclassical" theory of summation of more general sequences of random variables (see, for example, [4], Chap. 5, Sect. 6).

Now consider the class of characteristic functions $f(t)$ given by

$$F = \left\{ f(t) \,|\, f'(0) = 0, \; = -f''(0) = -\sigma^2 < \infty \right\}.$$

In the class F, we introduce the transformation (the Stein-Tikhomirov operator)

$$\Delta f(t) = f'(t) + t\sigma^2 f(t). \tag{6}$$

Obviously,

$$\Delta\left(e^{-t^2\sigma^2/2}\right) = 0, \tag{7}$$

i.e., the operator $\Delta(\cdot)$ "cancels" the normal characteristic function.

If we consider (6) as a differential equation to be solved for the initial condition $f(0) = 1$, then we obtain

$$f(t) - e^{-t^2\sigma^2/2} = e^{-t^2\sigma^2/2} \int_0^t \Delta(f(u))\, e^{u^2\sigma^2/2} du. \tag{8}$$

In relation (8), the sign of the variable of integration is identical with that of t and $|u| \leq |t|$. Relations (7) and (8) show that the expression $\Delta(f(t))$ characterizes the proximity of the distribution with characteristic function $f(t)$ to the normal law with mean 0 and variance σ^2.

It can be readily verified that the operator $\Delta(\cdot)$ possesses the following important property.

Lemma. *For characteristic functions $f(t)$ and $g(t)$ such that*

$$f'(0) = g'(0) = 0, \quad \max\left(\left|f''(0)\right|, \left|g''(0)\right|\right) < \infty$$

the following relation holds:

$$\Delta(f(t)\, g(t)) = f(t)\, \Delta(g(t)) + g(t)\, \Delta(f(t)). \tag{9}$$

It follows from this lemma that the operator $\Delta(\cdot)$ is the differentiation operator with respect to the product of characteristic functions.

Theorem 1. *In order that*

$$\sup_x |F_n(x) - \Phi(x)| \to 0$$

as $n \to \infty$, it is necessary and sufficient that for any $T > 0$ the following relation holds:

$$\sup_{|t|\le T} \sum_j \left|\Delta\left(f_{nj}(t)\right)\right| \to 0, \tag{10}$$

where $f_{nj}(\cdot)$ is the characteristic function corresponding to the distribution function $F_{nj}(x)$.

Proof. The proof of the sufficiency of condition (10) is simple enough. Indeed,

$$f_n(t) = Ee^{itS_n} = \prod_j f_{nj}(t)$$

and from relation (8) it follows that

$$\sup_{|t|\le T} \left|f_n(t) - e^{-t^2/2}\right| \le T \cdot \sup_{|t|\le T} |\Delta(f_n(t))| \tag{11}$$

for any $T > 0$.

Further, by (9) we have

$$\Delta(f_n(t)) = \sum_j \prod_{k\le j-1} f_{nk}(t)\Delta\left(f_{nj}(t)\right) \prod_{s\ge j+1} f_{ns}(t)$$

and, therefore,

$$|\Delta(f_n(t))| \le \sum_j \left|\Delta\left(f_{nj}(t)\right)\right|. \tag{12}$$

Relations (11) and (12) prove the necessity of condition (10) for the validity of the central limit theorem. To demonstrate the necessity of condition (10), let us prove that is not stronger than (5). Formally, this is sufficient, and the subsequent arguments will supply the necessary details. Set

$$\varphi_{nj}(t) = \int_{-\infty}^{\infty} e^{itx} d\Phi_{nj}(x) = \int_{-\infty}^{\infty} e^{itx} d\Phi\left(\frac{x}{\sigma_{nj}}\right).$$

Taking into account the fact that $\Delta\left(\varphi_{nj}(t)\right)=0$ for any $j\geq 1$, we have

$$\sum_j \left|\Delta\left(f_{nj}(t)\right)\right| = \sum_j \left|\Delta\left(f_{nj}(t)\right)-\Delta\left(\varphi_{nj}(t)\right)\right| \leq \sum_j \left|f'_{nj}(t)-\varphi'_{nj}(t)\right|$$

$$+|t|\sum_j \sigma_{nj}^2\left|f_{nj}(t)-\varphi_{nj}(t)\right| = \sum{}_1(t)+|t|\sum{}_2(t). \tag{13}$$

Noting that

$$EX_{nj}=0, \qquad \int_{-\infty}^{\infty} x^2 dF_{nj} = \int_{-\infty}^{\infty} x^2 d\Phi_{nj} = \sigma_{nj}^2,$$

and integrating by parts, we obtain

$$\left|f'_{nj}(t)-\varphi'_{nj}(t)\right| = \left|\int_{-\infty}^{\infty}(ix)\left(e^{itx}-1-itx\right)d\left(F_{nj}-\Phi_{nj}\right)\right|$$

$$\leq \left|\int_{-\infty}^{\infty}\left(e^{itx}-1-itx\right)\left(F_{nj}(x)-\Phi_{nj}(x)\right)dx\right|$$

$$+|t|\left|\int_{-\infty}^{\infty}(ix)\left(e^{itx}-1\right)\left(F_{nj}(x)-\Phi_{nj}(x)\right)dx\right|.$$

Therefore

$$\sum{}_1(t)\leq t^2\varepsilon\sum_i \int_{|x|\leq\varepsilon}|x|\left|F_{nj}(x)-\Phi_{nj}(x)\right|dx$$

$$+\left(|t|+t^2\right)\sum_i \int_{|x|\leq\varepsilon}|x|\left|F_{nj}(x)-\Phi_{nj}(x)\right|dx$$

$$\leq t^2\varepsilon\sum_i 2\sigma_{nj}^2+\left(|t|+t^2\right)R_n(\varepsilon)\leq 2\left(|t|+t^2\right)\left(\varepsilon+R_n(\varepsilon)\right). \tag{14}$$

To derive (14), we use the following fact. If $F(x)$ is a distribution function with mean 0 and variance σ^2, then

$$\int_0^{\infty} u\left(1-F(u)+F(-u)\right)du=\frac{\sigma^2}{2}.$$

It was established in [3] that

$$\sum\nolimits_{2}(t) \le 2\left(t^2 + |t|^3\right)(\varepsilon + R_n(\varepsilon)). \tag{15}$$

It follows from relations (13)–(15) that if condition (5) is satisfied, then for any $T > 0$ we have

$$\sup_{t \le T} \sum_j \left|\Delta\left(f_{nj}(t)\right)\right| \to 0, \quad n \to \infty.$$

We can easily verify condition (10) using the following simple example of increasing sums of independent Bernoulli random variables as an illustration. Suppose that

$$Y_j = \begin{cases} 1 & \text{with} \quad \text{probability} \quad p_j, \\ 0 & \text{with} \quad \text{probability} \quad q_j = 1 - p_j. \end{cases}$$

Taking into account the fact that $MY_j = p_j$, $DX_j = p_j q_j$, we set

$$B_n^2 = \sum_{j=1}^{n} p_j q_j, \quad X_{nj} = \frac{Y_j - p_j}{B_n}, \quad S_n = \sum_{j=1}^{n} X_{nj}.$$

In the case considered, we have

$$f_{nj}(t) = E^{itX_{nj}} = p_j e^{itq_j/B_n} + q_j e^{-itp_j/B_n}.$$

Let us show that if $B_n \to \infty$, then condition (10) holds. Indeed, it is easy to see that

$$f_{nj}(t) = 1 - \frac{p_j q_j}{2B_n^2} t^2 + \frac{p_j q_j}{B_n^2} \varepsilon_n(t), \tag{16}$$

$$f'_{nj}(t) = -\frac{p_j q_j}{B_n^2} t + \frac{p_j q_j}{B_n^2} \varepsilon'_n(t), \tag{17}$$

where

$$\sup_{|t| \le T} |\varepsilon_n(t)| = O\left(\frac{1}{B_n}\right), \quad n \to \infty,$$

for any $T > 0$.

It follows from relation (16) and (17) that, as $n \to \infty$, we have

$$\sup_{|t| \le T} \sum_j \left|\Delta(f_{nj}(t))\right| = O\left(\frac{1}{B_n}\right). \tag{18}$$

Obviously, for our sequence of simple random variables the direct verification of (5) or of the classical Lindeberg condition is more complicated that the estimates (18) obtained in this paper.

Remark 1. One can give more complicated examples of sequences of random variables for which the proof of the validity of the central limit theorem simplifies if the criterion (10) is used. Apparently, the present paper is the first paper in which the criterion for the convergtnce of the distribution of the sum S_n to the normal law is stated in terms of characteristic function of the summands.

Remark 2. In proving limit theorems for the distribution functions of sums of independent and weakly dependent random variables by the method of characteristic functions, on is mainly occupied with proving the fact that the characteristic function of these sums $f_n(t)$ does not vanish in a sufficiently large of neighborhood of the point $t = 0$. But there is no need for such a proof if we use the Stein-Tikhomirov method, this shows the advantage of this method over others.

Remark 3. Relation (8) and (11) show that the arguments used in the proof of the Theorem 1 allow us to obtain an estimate of the rate of convergence in the nonclassical case. Subsequent papers by this author will be concerned with exact statements and proofs for the corresponding assertion.

2 Berry-Esseen Inequality for Sampling Sums from Finite Population

Let $\{X_1, X_2, \ldots, X_N\}$ be a population of independent random variables and S_n be a sampling sum of size n. The last means that the sum S_n consist from such random variables which hit in a sample of size n from the parent population. One can give the exact meaning to the formation of the sum S_n as follows. Let $I = (I_1, I_2, \ldots, I_N)$ be an indicator random vectors such that $I_k = 0$ or $1 \quad (1 \le k \le N)$ and S_n contains the term X_k if and only if $I_k = 1$. Hence,

$$S_n = \sum_{k=1}^{N} I_k X_k.$$

It is assumed that I is independent from random variables $X_1, X_2, \ldots, X_N$ and for every ordered sequence $i = (i_1, i_2, \ldots, i_N)$ of n units and $N - n$ zeros

$$P\,(I = i) = \frac{1}{\binom{N}{n}} = \left(C_N^n\right)^{-1}.$$

We have $EI_k = \frac{n}{N} = f$ - the sampling ratio, and $EI_kI_i = \frac{n}{N} \cdot \frac{n-1}{N-1}$ for $k \neq i$. We introduce the moments $EX_k = \mu_k$, $EX_k^2 = \beta_k$ and then get

$$ES_n = \sum_{k=1}^{N} EI_kX_k = f\sum_{k=1}^{N}\mu_k,$$

$$ES_n^2 = \frac{n}{N}\sum_{k=1}^{N}\beta_k + \frac{n}{N}\cdot\frac{n-1}{N-1}\sum_{k\neq i}\mu_k\mu_i.$$

We will assume that (without loss of generality) the parent population of random variables has 0 mean and unit variance, i.e.

$$\sum_{k=1}^{N}\mu_k = 0, \quad \frac{1}{N}\sum_{k=1}^{N}\beta_k = 1. \tag{19}$$

Thus,

$$ES_n = 0, \quad DS_n = \mathrm{var}S_n = n\left(1 - \frac{n-1}{N-1}\alpha^2\right), \quad \alpha^2 = \frac{1}{N}\sum_{k=1}^{N}\mu_k^2.$$

We prove that $S_n/\sqrt{n}$ has approximately normal distribution with 0 mean and variance $1 - f\alpha^2$, and also give an estimation of the remainder term. In addition, the obtained result is a generalization of the classical Berry-Esseen estimation in CLT (S_n is turned into usual sum of n independent random variables when $n = N$).

The special case $X_i = a_i = const$ is very important in statistical applications of sampling sums. This case was investigated in details be B. Rosen [6]. The convergence rate in CLT were studied by A. Bikelis [1] in the case $X_i = const$ and by B. von Bahr [3] for arbitrary population of independent random variables. In the present work the result of last paper is made more precise.

Set

$$E|X_k|^3 = \gamma_k, \quad L_N = \frac{1}{N}\sum_{k=1}^{N}\gamma_k, \quad \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-u^2/2}du.$$

Theorem 2. *There exists an absolute positive constant C such that*

$$\sup_x \left| P\left(\frac{S_n}{\sqrt{n(1-f\alpha^2)}} < x\right) - \Phi(x)\right| \leq \frac{C\cdot L_N}{\sqrt{n}(1-f\alpha^2)^{3/2}}.$$

Remark 4. In [12] $C = 60$ and it is involved less exact characteristic

$$\gamma = \max_{1\leq k\leq N}\gamma_k$$

instead of L_N.

Remark 5. Rather rough calculation shows that $C < 60$ in given theorem, but we note that the exact calculation of the constant C doesn't enter is our task.

Remark 6. If the set of random variables $(X_1, X_2, \ldots, X_N)$ doesn't satisfy the normalizing conditions (19), we can easily obtain a new set $\left(X_1', X_2', \ldots, X_N'\right)$ which satisfies (1), by a linear transformation. Application of the result of Theorem 2 to this new set of random variables gives, in terms of the original variables

$$\left| P\left(\frac{S_n - n\mu}{\sqrt{\frac{1}{n}\left[\frac{1}{N}\sum\limits_{k=1}^{N}\sigma_k^2 + \frac{1-f}{N}\sum\limits_{k=1}^{N}(\mu_k-\mu)^2\right]}} < x \right) - \Phi(x) \right| \leq$$

$$\leq \frac{C}{\sqrt{n}} \cdot \frac{L_N}{\left[\frac{1}{N}\sum\limits_{k=1}^{N}\sigma_k^2 + \frac{1-f}{N}\sum\limits_{k=1}^{N}(\mu_k-\mu)^2\right]^{3/2}},$$

where

$$\mu_k = EX_k, \quad \mu = \frac{1}{N}\sum_{k=1}^{N}\mu_k \quad \text{and} \quad \sigma_k^2 = \text{var}X_k.$$

Proof of the Theorem 2 is conducted by means of the Stein-Tikhomirov method above mentioned at the point 1. Notice that in the papers [2, 8] are demonstrated application of initial variant of Stein-Tikhomirov method for obtained of classical Berry-Esseen inequality in the case of usual sum from independent random variables (i.e. as $(N = n)$). Let ν be a random variable with uniform distribution on the set $\{1, 2, \ldots, N\}$ that is not independent neither from random variables $X_1, X_2, \ldots, X_N$ nor from indicator vector I and F_{IX} be a σ-algebra generated by random variables$\{I_1, I_2, \ldots, I_N, X_1, X_2, \ldots, X_N\}$.

Further we denote

$$\omega_n = \frac{N}{\sqrt{DS_n}} I_\nu X_\nu.$$

It is not difficult to see that

$$\bar{S}_n = E\left(\omega_n / F_{IX}\right) = \frac{S_n}{\sqrt{DS_n}}. \tag{20}$$

Set also

$$f_n(t) = Ee^{it\bar{S}_n}.$$

As it follows from the point 1, we must calculate the operator $\Delta\left(f_n\left(t\right)\right)$ by the formula (6).

By virtue of (20)

$$E\left(i\omega_n e^{it\bar{S}_n}\right) = E\left[E\left(i\omega_n e^{it\bar{S}_n}/F_{IX}\right)\right] = E\left[iE\left(\omega_n/F_{IX}\right)e^{it\bar{S}_n}\right] = E\left(i\bar{S}_n e^{it\bar{S}_n}\right).$$

Therefore,

$$f_n'(t) = E\left(i\bar{S}_n e^{it\bar{S}_n}\right) = E\left(i\omega_n e^{it\bar{S}_n}\right). \tag{21}$$

By direct calculation we can obtain the following equalities:

$$E\omega_n = E\left(E\left(\omega_n / F_{IX}\right)\right) = E\bar{S}_n = 0. \tag{22}$$

$$E\omega_n^2 = \frac{N}{1-\frac{n-1}{N-1}\alpha^2} = \frac{n}{f\left(1-\frac{n-1}{N-1}\alpha^2\right)}, \tag{23}$$

$$E\left|\omega_n^3\right| = \frac{N^2}{\sqrt{n}\left(1-\frac{n-1}{N-1}\alpha^2\right)^{3/2}} \cdot L_N = \frac{n^2}{f^2 \cdot \sqrt{n}\left(1-\frac{n-1}{N-1}\alpha^2\right)^{3/2}}, \tag{24}$$

Further, set

$$S_{n\nu} = \frac{1}{\sqrt{DS_n}} \sum_{i\neq\nu} I_i X_i.$$

By virtue of (21) we have

$$f_n'(t) = E\left(i\omega_n e^{itS_{n\nu}}\right) + E\left[i\omega_n\left(e^{it\bar{S}_n} - e^{itS_{n\nu}}\right)\right].$$

Since ω_n and $S_{n\nu}$ are independent on construction, we have

$$E\left(i\omega_n e^{itS_{n\nu}}\right) = E\left(e^{itS_{n\nu}}\right) E\left(i\omega_n\right) = 0.$$

Thus,

$$f_n'(t) = E\left[i\omega_n\left(e^{it\omega_n/N} - 1\right)\right] \cdot Ee^{itS_{n\nu}}. \tag{25}$$

In addition

$$Ee^{itS_{n\nu}} = Ee^{it\bar{S}_n} + E\left(e^{itS_{n\nu}} - e^{it\bar{S}_n}\right) = f_n(t) + E\left[e^{itS_{n\nu}}\left(1-e^{it\omega_n/N}\right)\right]. \tag{26}$$

It follows from (25) and (26) that

$$f_n'(t) = E\left[i\omega_n\left(e^{it\omega_n/N}-1\right)\right] f_n(t) + E\left[i\omega_n\left(e^{it\omega_n/N}-1\right)\right] E\left(1-e^{it\omega_n/N}\right) Ee^{itS_{n\nu}}. \tag{27}$$

Using the equalities (22)–(24) we can obtain the following estimates

$$\left|E\left[i\omega_n\left(e^{it\omega_n/N}-1\right)\right] + t\right| \le \frac{t^2}{2}\frac{L_N}{\sqrt{n}(1-f\alpha^2)^{3/2}}, \tag{28}$$

$$\left|E\left(1-e^{it\omega_n/N}\right)\right| \le c_0 t^2 \frac{L_N}{\sqrt{n}(1-f\alpha^2)^{3/2}}, \tag{29}$$

In what follows, the letter c_0 denotes different absolute constants.

Now, with regard to the inequalities (28) and (29), we can rewrite (27) in the form

$$f_n'(t) = A_n(t) f_n(t) + B_n(t) \tag{30}$$

where

$$A_n(t) = -t + \frac{\theta}{2} t^2 \bar{L}_N, \quad |B_n(t)| \le c_0 t^2 \left| f_{n\nu}(t) \right| \bar{L}_N,$$

$$f_{n\nu}(t) = E e^{itS_{n\nu}}, \quad |\theta| \le 1, \quad \bar{L}_N = \frac{L_N}{\sqrt{n}\,(1 - f\alpha^2)^{3/2}}.$$

We can consider the equality (14) as the differential equation that we must to solve under the initial condition $f_n(0) = 1$. Then we have

$$f_n(t) = \exp\left\{ \int\limits_0^t A_n(u) du \right\} + \int\limits_0^t B_n(u) \exp\left\{ \int\limits_u^t A_n(s) ds \right\} du. \tag{31}$$

Further, we obtain

$$\int\limits_0^t A_n(u) du = -\frac{t^2}{2} + \frac{\theta}{6} \bar{L}_N \left| t \right|^3, \tag{32}$$

$$\int\limits_u^t A_n(s) ds = -\frac{t^2}{2} + \frac{u^2}{2} + a_n\left(t, u\right), \tag{33}$$

where

$$\left| a_n\left(t, u\right) \right| = \left| \theta \frac{\bar{L}_N}{2} \int\limits_u^t s^2 ds \right| \le \frac{\bar{L}_N}{2} \left| t \right| \left(t^2 - u^2 \right). \tag{34}$$

By direct calculation we obtain that

$$f_{n\nu}(t) = \sum_{j=1}^{N} E\left(e^{itS_{n\,\nu}}, \nu = j \right) = \frac{1}{N} \frac{1}{C_N^n} \sum_{j=1}^{N} \sum_{(r_1,\ldots,r_n)} \prod_{k=1}^{n}{}^{(j)} f_{r_k}\left(\frac{t}{\sqrt{DS_n}} \right), \tag{35}$$

where $f_j\left(t\right) = E e^{itX_j}$, $\prod^{(j)}$ means that in product $\prod\limits_{k=1}^{n} f_k(t)$ the factor with index r_j is equal to 1 and the summation is produced on all samples $(r_1, \ldots, r_n)$ of size n.

By using the paper [12] and (35) we can prove that under $|t| \le \left(\bar{L}_N\right)^{-1/3}$

$$\left| f_{n\nu}\left(t\right) \right| \le e^{-t^2/3}. \tag{36}$$

From (31)–(34), (36) we obtain finally that under $|t| \leq \left(\bar{L}_N\right)^{-1/3}$

$$\left| f_n(t) - e^{-t^2/2} \right| \leq c_0 \bar{L}_N \, |t|^3 \, e^{-t^2/6}. \tag{37}$$

Further way of the proof is the same as the proof of the classical Berry-Esseen inequality for sums from independent random variables (see [5]).

Acknowledgements The authors are thankful to an unknown referee for very helpful suggestions and also to professor A.N. Tikhomirov for submitted possibility of acquaintance with the paper [11].

References

1. Bikelis, A.: On the estimation of the remainder term in the central limit theorem for samples from finite populations. Studia Sci. Math. Hungar. **4**, 345–354 (1969)
2. Bolthausen, E.: An estimate of the remainder in a combinatorial limit theorem. Z. Wahr. Verw. Geb. **66**, 379–386 (1984)
3. Linnik, Yu.I., Ostrovskii, I.V.: Expansions of Random Variables and Vectors [in Russian]. Nauka, Moscow (1972)
4. Liptser, R.Sh., Shiryaev, A.N.: Martingale Theory [in Russian]. Nauka, Moscow (1986)
5. Petrov, V.: Limit Theorems for Sums of Independent Random Variables [in Russian]. Nauka, Moscow (1987)
6. Rosen, B. Limit theorems for sampling from finite populations. Arch. Math. **5**, 383–424 (1964)
7. Rotar', V.I.: On the generalization of the Lindeberg-Feller theorem. Mat. Zametki [Math. Notes], **18**(1), 129–135 (1975)
8. Soo-Thong, H., Chen L.H.Y. An L_p bound for the remainder in combinatorial central limit theorem. Ann. Probab. **6**(2), 231–249 (1978)
9. Stein, Ch.: A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In: Proceedings of the Sixth Berkley Symposium on Mathematical Statistics and Probability 2, pp. 583–602. University of California Press, Berkeley (1972)
10. Tikhornirov, A.N.: On the convergence of the remainder in the central limit theorem for weakly dependent random variables. Teor. Veroyatnost. i Primenen [Theory Probab. Appl.] **25**(4), 800–818 (1980)
11. Tikhornirov, A.N.: The rate of convergence in the central limit theorem for weakly dependent random variables. Vestnik SGU Ser. 1, **2**, 91–112 (1996)
12. von Bahr, B.: On sampling from a finite set of independent random variables. Z. Wahr. Verw. Geb. **24**, 279–286 (1972)