THEORY PROBAB. APPL.
Vol. 66, No. 4, pp. 537–549

THE CHEBYSHEV–EDGEWORTH CORRECTION IN THE CENTRAL LIMIT THEOREM FOR INTEGER-VALUED INDEPENDENT SUMMANDS*

S. G. BOBKOV† AND V. V. ULYANOV‡

(*Translated by the authors*)

Abstract. We give a short overview of the results related to the refined forms of the central limit theorem, with a focus on independent integer-valued random variables (r.v.'s). In the independent and non-identically distributed (non-i.i.d.) case, an approximation is then developed for the distribution of the sum by means of the Chebyshev–Edgeworth correction containing the moments of the third order.

Key words. central limit theorem, the Chebyshev–Edgeworth correction, integer-valued random variables

DOI. 10.1137/S0040585X97T990605

1. Introduction.

Let $X_1, \dots, X_n$ be independent random variables (r.v.'s) with finite absolute moments of the third order. Consider the sum $S_n = X_1 + \cdots + X_n$. It is known that S_n has a nearly normal distribution with mean $\mu = \mathbf{E} S_n$ and variance $\sigma^2 = \mathbf{D} S_n$ $(\sigma > 0)$, as long as the third-order Lyapunov ratio (or fraction)

$$L_3 = \frac{1}{\sigma^3} \sum_{k=1}^{n} \mathbf{E}|X_k - \mathbf{E}X_k|^3$$

is small. A quantitative result is given by the Berry–Esseen inequality

$$\sup_x \left| \mathbf{P}\{S_n \leqslant x\} - \Phi\left(\frac{x-\mu}{\sigma}\right) \right| \leqslant c L_3, \tag{1.1}$$

holding with some positive absolute constant c (cf., e.g., [19]). Here and in what follows, Φ denotes the standard normal distribution function (d.f.) with probability density function

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}, \qquad x \in \mathbf{R}.$$

We necessarily have $L_3 \geqslant 1/\sqrt{n}$. In the i.i.d. case $X_k = \xi_k/\sqrt{n}$, this inequality can be reversed up to a factor depending on ξ_1. Hence in (1.1) we have

$$\mathbf{P}\{S_n \leqslant x\} = \Phi\left(\frac{x-\mu}{\sigma}\right) + O\left(\frac{1}{\sqrt{n}}\right), \qquad n \to \infty, \tag{1.2}$$

with a standard rate of normal approximation.

*Received by the editors May 31, 2021. This research was carried out within the framework of the Moscow Center for Fundamental and Applied Mathematics, Lomonosov Moscow State University, and HSE University Basic Research Programs. The paper was published with the financial support of the Ministry of Education and Science of the Russian Federation as part of the program of the Moscow Center for Fundamental and Applied Mathematics under the agreement 075-15-2019-1621. The research of S. Bobkov was partially supported by NSF grant DMS-1855575. Originally published in the Russian journal *Teoriya Veroyatnostei i ee Primeneniya*, 66 (2021), pp. 676–692.

https://doi.org/10.1137/S0040585X97T990605

†School of Mathematics, University of Minnesota, Minneapolis, MN 55455 USA (bobkov@math.umn.edu).

‡Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow, Russia, and National Research University Higher School of Economics, Moscow, Russia (vulyanov@cs.msu.ru).

In order to make a more precise statement with a smaller error of approximation than (1.1), (1.2), the normal distribution function should be slightly corrected in a smooth way. Namely, we introduce

$$\Phi_3(x) = \Phi(x) - \frac{l_3}{6}(x^2-1)\varphi(x), \qquad x \in \mathbf{R}, \tag{1.3}$$

where

$$l_3 = \frac{1}{\sigma^3}\,\mathbf{E}(S_n - \mathbf{E}S_n)^3 = \frac{1}{\sigma^3}\sum_{k=1}^{n}\mathbf{E}(X_k - \mathbf{E}X_k)^3.$$

Here the index 3 reflects the fact that in (1.3) the moments of X_k up to the third order are used. In view of the Fourier–Stieltjes transform of a signed Borel measure on the real line, the function Φ_3 is designated below as the third-order Chebyshev–Edgeworth correction of the d.f. $F_n = \mathbf{P}\{Z_n \leqslant x\}$ of the normalized sum $Z_n = (S_n - \mu)/\sigma$. In the i.i.d. case with $X_k = \xi_k/\sqrt{n}$, the correction $\Phi_3(x) - \Phi(x)$ appears as the first term in the expansion for $F_n(x)$ in powers of $1/\sqrt{n}$ in the form

$$F_n(x) - \Phi(x) \sim Q_1(x)\varphi(x)\frac{1}{n^{1/2}} + Q_2(x)\varphi(x)\frac{1}{n} + \cdots .$$

Here, each $Q_j(x)$ represents a polynomial whose coefficients depend on the first $j+2$ moments of ξ_1. Based on the idea of expansion of arbitrary functions in series of Chebyshev–Hermite polynomials, the study of such expansions was started in 1887 by Chebyshev [7] and then continued by Edgeworth, Charlier, Cramér, and Esseen, among others. For references and discussion of the subject, we refer the reader to Gnedenko and Kolmogorov [15].

Even if we restrict ourselves to the first term in this expansion, a general problem is to explore whether or not it is possible to improve inequality (1.1) by replacing Φ with Φ_3. Such a replacement would not deteriorate this bound in view of the relation $|l_3| \leqslant L_3$. On the other hand, a comparison of smooth linear functionals (for example, characteristic functions (ch.f.'s)) of F_n and Φ_3 suggests that an improvement is indeed possible in various natural scenarios.

2. Nonlattice distributions. In particular, in the i.i.d. situation, assuming that X_1 has a nonlattice distribution, Esseen [12] derived the representation

$$\mathbf{P}\{S_n \leqslant x\} = \Phi_3\biggl(\frac{x-\mu}{\sigma}\biggr) + o\biggl(\frac{1}{\sqrt{n}}\biggr), \qquad n \to \infty, \tag{2.1}$$

which holds uniformly over all $x \in \mathbf{R}$, where as before, $\mu = n\,\mathbf{E}X_1$ and $\sigma^2 = n\mathbf{D}(X_1)$. Although somewhat implicitly, this theorem improves the standard rate of normal approximation as in (1.2). The remainder term in (2.1) can be improved to $O(1/n)$, provided that $\mathbf{E}X_1^4 < \infty$ and assuming that the Cramér continuity condition

$$\limsup_{t\to\infty} |\mathbf{E}e^{itX_1}| < 1, \tag{2.2}$$

is fulfilled; this is a particular case of Cramér's theorem (see [9] and [10]).

In general, however, the order of magnitude of the remainder term depends on arithmetical properties of the point spectrum of the d.f. of X_1. If condition (2.2) is not met, the Kolmogorov distance

$$\Delta_n = \sup_x \biggl|\mathbf{P}\{S_n \leqslant x\} - \Phi_3\biggl(\frac{x-\mu}{\sigma}\biggr)\biggr|$$

may actually be of the order of $n^{-\alpha}$ up to logarithmic factors, for any prescribed value α, $\frac{1}{2} < \alpha < 1$. Let us mention the following characterization for the i.i.d. case $X_k = \xi_k/\sqrt{n}$ with $\mathbf{E}\xi_1^4 < \infty$ in terms of the (common) ch.f. $f(t) = \mathbf{E}\, e^{it\xi_1}$. Namely, given $p \geqslant 2$, the property

$$\Delta_n = \widetilde{O}(n^{1/2-1/p}), \qquad n \to \infty, \tag{2.3}$$

is equivalent to saying that

$$\frac{1}{1-|f(t)|} = \widetilde{O}(t^p), \qquad t \to \infty. \tag{2.4}$$

Here, we use the notation $\widetilde{O}(t^p)$ for the growth rate $O(t^p\,(\ln t)^q)$ with some $q \in \mathbf{R}$ and similarly $\widetilde{O}(n^p)$ for $O(n^p(\ln n)^q)$.

Simple discrete examples, where ξ_1 takes 4 values ± 1, $\pm a$, each with probability $\frac{1}{4}$ for irrational numbers a, are described in [6] (note that $\Phi_3 = \Phi$ for symmetric distributions).

Let

$$\eta(a) = \sup\Big\{\eta > 0\colon\ \liminf_{n\to\infty} n^\eta \|na\| = 0\Big\} = \inf\Big\{\eta > 0\colon\ \inf_{n\geqslant 1} n^\eta \|na\| > 0\Big\},$$

where $\|x\|$ denotes the distance from a real number x to the closest integer. The value $\eta = \eta(a)$, called a type of an irrational number a, is optimal in the sense that, for any $\varepsilon > 0$, the Diophantine inequality $|a - p/q| < q^{-(1+\eta-\varepsilon)}$ has infinitely many rational solutions p/q. By Dirichlet's theorem, $\eta \geqslant 1$. The possible values of η fills the whole half-axis $[1, \infty]$ including the case $\eta = \infty$ (which describes the Liouville numbers); cf. [1]. Applying the equivalence of (2.3) and (2.4) with $p = 2\eta$, we have, for any $\varepsilon > 0$,

$$\Delta_n = O(n^{-1/2-1/(2\eta)+\varepsilon}), \qquad n \to \infty,$$

if and only if the number a is of type η.

Note that the multidimensional case differs markedly from the one-dimensional case. For instance, the distribution of the normalized sum of i.i.d. random vectors is approximated only by a Gaussian distribution without any corrections on the class of all centered ellipsoids with an accuracy of the order from $o(1/\sqrt{n})$ up to $O(1/n)$. It holds under the appropriate dimension of space and when the summands satisfy some moment conditions, for example, finiteness of the fourth absolute moment (see, e.g., [12], [20], and [14]). For the non-i.i.d. random vectors case, see [23].

3. Lattice distributions. Bernoulli and Poisson schemes. It was also shown by Esseen that a representation similar to (2.1) holds also for lattice distributions if one adds to Φ_3 a certain discontinuous periodic function with a factor of the order of $1/\sqrt{n}$. To make the statement more transparent, we suppose without loss of generality that X_1 takes integer values, with $h = 1$ being the maximal step (span), so that the distribution of X_1 is not supported on $h\mathbf{Z}$ for $h > 1$. Now Theorem 3 in [12, p. 56] can be equivalently stated as

$$\mathbf{P}\{S_n \leqslant x\} = \Phi_3\bigg(\frac{x-\mu}{\sigma}\bigg) + \frac{1}{\sigma}\,\psi(x)\varphi\bigg(\frac{x-\mu}{\sigma}\bigg) + o\bigg(\frac{1}{\sqrt{n}}\bigg), \tag{3.1}$$

with $\psi(x) = x - [x] + \frac{1}{2}$ (the convergence is uniform in x; cf. also [15, Chap. 8, section 43]. This result was refined by Bikjalis [4] with a nonuniform remainder term and then by Osipov [17] to higher order Chebyshev–Edgeworth expansions (thereby refining Theorem 4 in [12]). However, the additional Esseen terms in such representations have a rather complicated structure and are expressed in the form of infinite Fourier series.

To avoid unnecessary technicalities, we consider the probabilities $\mathbf{P}\{S_n \leqslant k\}$ for integers k only, since the sum S_n in (3.1) is integer-valued. Now the additional term vanishes if (3.1) is applied with $x = k + \frac{1}{2}$, which gives us the simpler representation

$$\mathbf{P}\{S_n \leqslant k\} = \Phi_3\biggl(\frac{k+1/2-\mu}{\sigma}\biggr) + o\biggl(\frac{1}{\sqrt{n}}\biggr), \tag{3.2}$$

which holds uniformly over all $k \in \mathbf{Z}$. As we see, the points where Φ_3 is evaluated in the two scenarios in (2.1) and (3.2) are slightly different. This well-known phenomenon should not be confusing; it was a focus of many investigations in the scheme of Bernoulli trials including the works by Bernstein [2], [3], Feller [13], and Uspensky [24]. If X_k takes only two values 1 and 0 with probabilities p and $q = 1-p$, we have $\mu = np$, $\sigma^2 = npq$, and now inequality (1.1) becomes

$$\sup_{0\leqslant k\leqslant n}\biggl|\mathbf{P}\{S_n \leqslant k\} - \Phi\biggl(\frac{k-np}{\sqrt{npq}}\biggr)\biggr| \leqslant \frac{c}{\sqrt{npq}}.$$

In his book [24, pp. 129–131], Uspensky established a two-term approximation implying the much stronger inequality

$$\sup_{0\leqslant k\leqslant n}\biggl|\mathbf{P}\{S_n \leqslant k\} - \Phi_3\biggl(\frac{k+1/2-np}{\sqrt{npq}}\biggr)\biggr| \leqslant \frac{c}{npq}, \tag{3.3}$$

which also quantifies the remainder term in (3.2). Here, according to (1.3), the Chebyshev–Edgeworth correction may be simplified to read

$$\Phi_3(x) = \Phi(x) - \frac{p-q}{6\sqrt{npq}}(x^2-1)\,\varphi(x).$$

Uspenksy's approach was adapted by Cheng [8] to get a Poissonian analogue of bound (3.3). It was shown that, if an r.v. ξ has a Poisson distribution with parameter $\lambda > 0$, that is,

$$\mathbf{P}\{\xi = k\} = \frac{\lambda^k}{k!}\,e^{-\lambda}, \qquad k = 0, 1, \dots,$$

then

$$\sup_{k}\biggl|\mathbf{P}\{\xi \leqslant k\} - \Phi_3\biggl(\frac{k+1/2-\lambda}{\sqrt{\lambda}}\biggr)\biggr| \leqslant \frac{c}{\lambda}. \tag{3.4}$$

Here, the Chebyshev–Edgeworth correction for ξ is given by

$$\Phi_3(x) = \Phi(x) - \frac{1}{6\sqrt{\lambda}}(x^2-1)\varphi(x), \qquad x \in \mathbf{R}, \tag{3.5}$$

(which does not depend on n). This representation is consistent with (1.3): it suffices to represent ξ as the sum of n independent Poisson r.v.'s with parameter λ/n.

The next natural step was made by Deheuvels, Puri, and Ralescu [11] who extended inequality (3.3) to independent Bernoulli r.v.'s X_k taking the values 0 and 1 with not necessarily equal probabilities $p_k = \mathbf{P}\{X_k = 1\}$. Namely, we similarly have

$$\sup_{0\leqslant k\leqslant n}\biggl|\mathbf{P}\{S_n \leqslant k\} - \Phi_3\biggl(\frac{k+1/2-\mu}{\sigma}\biggr)\biggr| \leqslant \frac{c}{\sigma^2} \tag{3.6}$$

for some absolute constant $c > 0$, where

$$\mu = p_1 + \cdots + p_n, \quad \sigma^2 = p_1 q_1 + \cdots + p_n q_n \quad (q_k = 1 - p_k),$$

and where Φ_3 is defined according to (1.3) with

$$l_3 = \frac{1}{\sigma^3} \sum_{k=1}^{n} p_k q_k (p_k - q_k).$$

For statistical reasons, this estimate was polished by Mikhailov [16], who showed that the right-hand side in (3.6) can be replaced by $(\sigma + 3)/(4\sigma^3)$ provided that $\sigma \geqslant 10$. See also [25] for a further improvement.

4. Further developments. The aim of the remaining part of this note is to extend estimate (3.6) to general independent integer-valued r.v.'s under the fourth moment condition (which would also contain the Poissonian case (3.4)). To this aim, we involve the Lyapunov ratio of order 4 defined by

$$L_4 = \frac{1}{\sigma^4} \sum_{k=1}^{n} \mathbf{E}(X_k - \mathbf{E}X_k)^4, \qquad \sigma^2 = \mathbf{D}S_n = \sum_{k=1}^{n} \mathbf{D}X_k. \tag{4.1}$$

This functional often appears naturally in various asymptotic expansions related to the central limit theorem. However, for our purposes this functional is insufficient, and so we introduce another quantity not related to the moments.

DEFINITION 4.1. *Given an integer-valued r.v. ξ with ch.f. $v(t) = \mathbf{E}\, e^{it\xi}$, $t \in \mathbf{R}$, we put*

$$\mathrm{V}(\xi) = -\sup_{0<t<2\pi} \frac{\ln|v(t)|}{1 - \cos t}. \tag{4.2}$$

One important feature of this functional is described in the following.

PROPOSITION 4.1. *If ξ is an integer-valued r.v., then $0 \leqslant \mathrm{V}(\xi) < \infty$. Moreover, $\mathrm{V}(\xi) > 0$ if and only if the distribution of ξ is nondegenerate and has the maximal step $h = 1$.*

In some sense, $\mathrm{V}(\xi)$ quantifies the "strength" of the property that the maximal step of the lattice distribution of ξ is exactly $h = 1$. To illustrate this, suppose that $\mathbf{P}\{\xi = \pm 1\} = 1 - \varepsilon/2$ and $\mathbf{P}\{\xi = 0\} = \varepsilon$ for some $\varepsilon \in (0,1)$. Then $h = 1$, while $\mathrm{V}(\xi) \to 0$ as $\varepsilon \to 0$. It is therefore not surprising that the limit distribution has a larger maximal step $h = 2$.

If the r.v. ξ has a finite second moment, then by applying the Taylor formula to the function $|v(t)|^2$ near zero, we have by (4.2)

$$\mathrm{V}(\xi) \leqslant \mathbf{D}\xi. \tag{4.3}$$

However, in general, it is unnecessary for these functionals to be of the same order, as the previous example shows, where $\mathbf{D}(\xi) \to 1$ as $\varepsilon \to 0$. On the other hand, we have

$$\mathrm{V}(\xi) = \mathbf{D}\xi = pq,$$

when ξ has a Bernoulli distribution with parameters $p = \mathbf{P}\{\xi = 1\}$ and $q = \mathbf{P}\{\xi = 0\}$.

Let us also point out a superadditivity property of the functional V along convolutions. From Definition 4.1, for the sum $S_n = X_1 + \cdots + X_n$ of independent integer-valued r.v.'s, we have

$$\mathrm{V}(S_n) \geqslant \sum_{k=1}^{n} \mathrm{V}(X_k). \tag{4.4}$$

Moreover, here in the i.i.d. case we have an equality. In particular, $\mathrm{V}(S_n) = \mathbf{D}(S_n) = npq$ for the binomial distribution with parameters (n,p). More generally, the equality in (4.4) also holds for sums of non-i.i.d. Bernoulli r.v.'s, since the supremum in (4.2) with $\xi = S_n$ is attained asymptotically at $t = 0$.

We can now formulate the main result. Recall that the Chebyshev–Edgeworth correction $\Phi_3(x)$ is defined in (1.3) and $\mathrm{V}(\xi)$ is defined in (4.2).

THEOREM 4.1. *Let integer-valued r.v.'s $X_1, \dots, X_n$ be independent and have finite fourth moments. For the sum $S_n = X_1 + \cdots + X_n$, we put $\mu = \mathbf{E}S_n$, $\sigma^2 = \mathbf{D}(S_n)$, and $\mathrm{V} = \sum_{k=1}^{n} \mathrm{V}(X_k)$. Then*

$$\sup_{k \in \mathbf{Z}} \biggl| \mathbf{P}\{S_n \leqslant k\} - \Phi_3\biggl(\frac{k + 1/2 - \mu}{\sigma}\biggr)\biggr| \leqslant \frac{c\sigma^2}{\mathrm{V}}\, L_4 \tag{4.5}$$

with an absolute constant $c > 0$.

In the i.i.d. case, the right-hand side of (4.5) is simplified, and we arrive at (3.2) with an improved remainder term.

COROLLARY 4.1. *Suppose that integer-valued r.v.'s X_k are independent and have a common nondegenerate distribution with maximal step $h = 1$ and $\mathbf{E}X_1^4 < \infty$. Let $\mu = n\,\mathbf{E}X_1$ and $\sigma^2 = n\,\mathbf{D}(X_1)$. Then*

$$\mathbf{P}\{S_n \leqslant k\} = \Phi_3\biggl(\frac{k + 1/2 - \mu}{\sigma}\biggr) + O\biggl(\frac{1}{n}\biggr) \tag{4.6}$$

as $n \to \infty$ uniformly over all $k \in \mathbf{Z}$.

Moreover, the involved constant in the remainder term does not exceed, up to a numerical factor, the quantity

$$\frac{\mathbf{E}(X_1 - \mathbf{E}X_1)^4}{n\,\mathrm{V}(X_1)\mathbf{D}X_1}.$$

If X_k are independent Bernoulli r.v.'s with parameters $p_k = \mathbf{P}\{X_k = 1\}$ and $q_k = \mathbf{P}\{X_k = 0\}$, then

$$\mathbf{E}(X_k - \mathbf{E}X_k)^4 = p_k q_k (p_k^3 + q_k^3) \leqslant p_k q_k, \quad \mathbf{E}(X_k - \mathbf{E}X_k)^4 \geqslant \frac{1}{4} p_k q_k.$$

Therefore, according to (4.1),

$$\frac{1}{4\sigma^2} \leqslant L_4 \leqslant \frac{1}{\sigma^2}. \tag{4.7}$$

In addition, $\sum_{k=1}^{n} \mathrm{V}(X_k) = \sigma^2$, which is mentioned above. As a result, inequality (4.5) contains (3.6) as a particular case (the Deheuvels–Puri–Ralescu theorem).

Another important particular case worth mentioning is related to the normal approximation for the Poisson distribution. If an r.v. ξ has Poisson distribution with parameter $\lambda > 0$, then, for all $t \in \mathbf{R}$, its ch.f. $v(t)$ satisfies

$$\frac{\ln|v(t)|}{1-\cos t} = -\lambda.$$

Hence $\mathrm{V}(\xi) = \mathbf{D}\xi = \lambda$. We also have $\mathbf{E}\xi = \lambda$, so it is easy to check that

$$\mathbf{E}(\xi - \mathbf{E}\xi)^3 = \lambda, \qquad \mathbf{E}(\xi - \mathbf{E}\xi)^4 = \lambda(3\lambda + 1). \tag{4.8}$$

Representing $\xi = X_1 + \cdots + X_n$ with independent Poisson r.v.'s X_k with parameter λ/n, we get $l_3 = 1/\sqrt{\lambda}$, so, the Chebyshev–Edgeworth correction for ξ is given by (3.5). In addition, $L_4 = 3/n + 1/\lambda$, according to (4.1) and (4.8). Hence, making $n \to \infty$, we get Cheng's bound (3.4) from Theorem 4.1.

Let us now return to Proposition 4.1 and describe a simple argument in the proof, which is required only in one part of the proof. If the distribution of an r.v. ξ is nondegenerate, then $|v(t)| \leqslant e^{-ct^2}$ for some $c > 0$ in a sufficiently small interval $|t| \leqslant t_0$ ($t_0 > 0$). In particular,

$$\frac{\ln(1/|v(t)|)}{1-\cos t} \geqslant \frac{ct^2}{1-\cos t} \geqslant c$$

in some neighborhood of zero. If, in addition, one knows that ξ is integer-valued, then $v(t)$ is (2π)-periodic, implying that $|v(t)| \leqslant e^{-c(t-2\pi)^2}$ for $|t - 2\pi| \leqslant t_0$. Hence

$$\frac{\ln(1/|v(t)|)}{1-\cos t} \geqslant c\frac{(t-2\pi)^2}{1-\cos t} \geqslant c$$

in some neighborhood of 2π. Finally, recall that the property that the maximal step is equal to $h = 1$ is equivalent to saying that $|v(2\pi)| = 1$ with $|v(t)| < 1$ for all $0 < t < 2\pi$ (cf. [19, Chap. 1, Lemma 1.2]). By the continuity of $v(t)$, the ratio in (4.2) is therefore bounded away from zero on the whole interval $(0, 2\pi)$.

5. Preparation to the proof. Here we collect some technical results required for the proof of Theorem 4.1.

Given independent r.v.'s $X_1, \dots, X_n$ with finite fourth moments, we put $\mu_k = \mathbf{E}X_k$, $\mu = \mu_1 + \cdots + \mu_n$, $\sigma^2 = \mathbf{D}(S_n)$ ($\sigma > 0$), and define

$$\xi_k = \frac{X_k - \mu_k}{\sigma}, \qquad Z_n = \xi_1 + \cdots + \xi_n = \frac{S_n - \mu}{\sigma}.$$

Clearly, $\mathbf{E}\xi_k = \mathbf{E}Z_n = 0$ and $\mathbf{D}Z_n = 1$.

Note that the Lyapunov fractions

$$L_p = \frac{1}{\sigma^p}\sum_{k=1}^n \mathbf{E}|X_k - \mathbf{E}X_k|^p, \qquad p \geqslant 2,$$

are affine invariant functionals, and hence L_3 and L_4 for the collection $\xi_1, \dots, \xi_n$ are the same as for $X_1, \dots, X_n$. A similar result also holds for l_3, and hence the Chebyshev–Edgeworth correction $\Phi_3(x)$ for Z_n is again given by (1.3).

The proximity of the d.f.

$$F_n(x) = \mathbf{P}\{Z_n \leqslant x\}, \qquad x \in \mathbf{R},$$

to $\Phi_3(x)$ in a weak sense can be studied in terms of the proximity of the ch.f.

$$f_n(t) = \mathbf{E}e^{itZ_n} = \int_{-\infty}^{\infty} e^{itx}\, dF_n(x)$$

to the Fourier–Stieltjes transform of Φ_3, that is, to the corrected normal ch.f.

$$g(t) = \int_{-\infty}^{\infty} e^{itx}\, d\Phi_3(x) = e^{-t^2/2}\biggl(1 + \frac{l_3}{6}(it)^3\biggr), \qquad t \in \mathbf{R}. \tag{5.1}$$

In particular, we have

LEMMA 5.1. *On the interval* $|t| \leqslant 1/L_3$,

$$|f_n(t) - g(t)| \leqslant cL_4 \min(1, t^4)\, e^{-t^2/8} \tag{5.2}$$

with some absolute constant $c > 0$.

This result, including Chebyshev–Edgeworth expansions for products of ch.f.'s of higher order (especially in the i.i.d. case, cf., e.g., [15, Chap. 8, section 40] or [18, Chap. 6, section 3]), is well known. The formulation of inequalities such as (5.2) in the non-i.i.d. case is often different in different places with respect to the interval where the bound holds and to the constants in the exponent (cf. [10, Chap. 7], [21], [22]). Our formulation follows that of [5, Theorem 18.1]. It is important that (5.2) implies the integral estimate

$$\int_{|t|<1/L_3} \biggl|\frac{f_n(t) - g(t)}{t}\biggr|\, dt \leqslant cL_4. \tag{5.3}$$

Let us briefly comment on the relationship between different Lyapunov fractions. Since $L_2 = 1$, the function $p \to L_p^{1/p-2}$ is nondecreasing with respect to $p > 2$. In particular,

$$L_3 \leqslant \sqrt{L_4}. \tag{5.4}$$

For integer-valued r.v.'s X_k, there are other relations involving the variance σ^2. According to the well-known von Mises inequality,

$$\mathbf{E}|\xi - \mathbf{E}\xi|^p \leqslant 2\mathbf{E}|\xi - \mathbf{E}\xi|^{p+1}, \qquad p \geqslant 1,$$

provided that an r.v. ξ takes only integer values.

For the reader's convenience, we give a short proof of this inequality. Let η be an integer-valued r.v. with zero mean and with all finite moments. By Hölder's inequality, for arbitrary positive p and q, and θ, $0 \leqslant \theta \leqslant 1$, we have

$$\mathbf{E}|\eta|^{\theta p + (1-\theta)q} \leqslant (\mathbf{E}|\eta|^p)^\theta (\mathbf{E}|\eta|^q)^{(1-\theta)}.$$

Hence, the function $R(p) = \ln \mathbf{E}|\eta|^p$ is convex for $p > 0$, and therefore $R(p+1) - R(p)$ is a nondecreasing function for $p > 0$. So, for $p \geqslant 1$, we have

$$\frac{\mathbf{E}|\eta|^{p+1}}{\mathbf{E}|\eta|^p} \geqslant \frac{\mathbf{E}\eta^2}{\mathbf{E}|\eta|} \geqslant \frac{1}{2},$$

since, if ζ is an independent copy of η, we have

$$\mathbf{E}\eta^2 = \frac{1}{2}\mathbf{E}(\eta-\zeta)^2 \geqslant \frac{1}{2}\mathbf{E}|\eta-\zeta| \geqslant \frac{1}{2}\mathbf{E}|\eta|,$$

because ζ and η are integer-valued. Hence, by the von Mises inequality,

$$L_3 \geqslant \frac{1}{2\sigma}, \quad L_4 \geqslant \frac{1}{2\sigma}L_3 \geqslant \frac{1}{4\sigma^2}. \tag{5.5}$$

The inequalities in (5.5) can be reversed, up to constants, for Bernoulli distributions (see, e.g., (4.7)).

For the function in (5.1), the following integral estimate complements (5.3).

LEMMA 5.2. *The following inequality holds*:

$$\int_{|t|>1/L_3} \left|\frac{g(t)}{t}\right| dt \leqslant 3L_4. \tag{5.6}$$

Proof. Given $T>0$, we have

$$\int_T^\infty \frac{1}{t}e^{-t^2/2}\,dt < \frac{1}{T^2}e^{-T^2/2} < \frac{1}{T^2},$$
$$\int_T^\infty t^2 e^{-t^2/2}\,dt < Te^{-T^2/2} < \frac{1}{T}.$$

Applying these inequalities with $T = 1/L_3$ and using (5.1) and (5.4), we see that the integral in (5.6) is majorized by

$$2L_3^2 + \frac{1}{3}|l_3|L_3 < 3L_3^2 \leqslant 3L_4,$$

the result required.

Following [11] and especially [16], where many arguments in the proof were clarified, let us now describe a smoothing operation that allows one to properly modify the Fourier analysis of the distribution of $S_n = X_1 + \cdots + X_n$. Let η denote an r.v. which is independent of all X_k and has a uniform distribution U on the interval $(-1/2, 1/2)$. If all X_k are integer-valued, then the r.v. $\widetilde{S}_n = S_n + \eta$ has an absolutely continuous distribution satisfying

$$\mathbf{P}\left\{\widetilde{S}_n \leqslant k + \frac{1}{2}\right\} = \mathbf{P}\{S_n \leqslant k\}, \qquad k \in \mathbf{Z}. \tag{5.7}$$

The d.f. $\widetilde{F}_n$ of $\widetilde{S}_n$ has a simple structure: one may just restrict the d.f. of S_n to the lattice $\mathbf{Z}$ and then extend it to the whole real line as a continuous function which is linear on every interval $[k, k+1]$.

In view of (5.7), one may now focus on the asymptotic approximation for $\widetilde{F}_n$. This can be done by employing the Chebyshev–Edgeworth correction for the extended sequence $X_1, \dots, X_n, \eta$. In other words, using the Chebyshev–Edgeworth correction Φ_3 for Z_n as in (1.3), it makes sense to approximate the d.f. of the smoothed r.v.

$$\widetilde{Z}_n = Z_n + \frac{\eta}{\sigma} = \frac{S_n - \mu + \eta}{\sigma}$$

by the convolution

$$\widetilde{\Phi}_3(x) = \Phi_3(x) * U(\sigma x) = \int_{-1/2}^{-1/2} \Phi_3\Big(x - \frac{y}{\sigma}\Big)\,dy. \tag{5.8}$$

Similarly to (5.2) and (5.3), the difference between Φ_3 and $\widetilde{\Phi}_3$ is small, as long as L_4 is small.

LEMMA 5.3. *The following inequality holds*:

$$\sup_x \big|\widetilde{\Phi}_3(x) - \Phi_3(x)\big| \leqslant \frac{1}{96\sigma^2}(1 + L_3). \tag{5.9}$$

Proof. By Taylor's integral formula applied to the integrand in (5.8),

$$\widetilde{\Phi}_3(x) - \Phi_3(x) = \frac{1}{\sigma^2} \int_{|y|<1/2} \int_{0<t<1} (1-t) y^2 \Phi_3''\Big(x - \frac{ty}{\sigma}\Big)\,dy\,dt,$$

which implies

$$\sup_x \big|\widetilde{\Phi}_3(x) - \Phi_3(x)\big| \leqslant \frac{1}{24\sigma^2} \sup_x |\Phi_3''(x)|.$$

According to (1.3),

$$\Phi_3''(x) = -x\varphi(x) - \frac{l_3}{6} H_4(x)\varphi(x),$$

where $H_4(x) = x^4 - 6x^2 + 3$ is the fourth-order Chebyshev–Hermite polynomial. It can be easily checked that

$$|x|\varphi(x) \leqslant \frac{1}{\sqrt{2\pi e}} < \frac{1}{4}, \qquad |H_4(x)|\varphi(x) \leqslant \frac{3}{\sqrt{2\pi}} < 1.2.$$

Using $|l_3| \leqslant L_3$, we get

$$|\Phi_3''(x)| < \frac{1}{4} + \frac{1}{5} L_3$$

and therefore (5.9). The lemma is proved.

6. Proof of Theorem 4.1. Under the notation of the previous section, recall that $f_n(t)$ denotes the ch.f. of the normalized sum Z_n. Hence, the ch.f. of the sum $S_n = X_1 + \cdots + X_n = \mu + \sigma Z_n$ is given by

$$u_n(t) = e^{i\mu t} f_n(\sigma t), \qquad t \in \mathbf{R}.$$

Similarly, we introduce

$$u(t) = \int_{-\infty}^{\infty} e^{itx}\,d\Phi_3\Big(\frac{x-\mu}{\sigma}\Big) = e^{i\mu t} g(\sigma t),$$

where g represents the Fourier–Stieltjes transform of Φ_3; cf. (5.1). Since the ch.f. of the r.v. $\eta \sim U$ is given by

$$\omega(t) = \mathbf{E}e^{it\eta} = \frac{\sin(t/2)}{t/2},$$

the ch.f. and the Fourier–Stieltjes transform of $\widetilde{S}_n = S_n + \eta$ and $\widetilde{\Phi}_3(x - \mu/\sigma)$ are given, respectively, by

$$\widetilde{u}_n(t) = u_n(t)\omega(t), \quad \widetilde{u}(t) = u(t)\omega(t).$$

Recall that $\widetilde{F}_n(x) = \mathbf{P}\{\widetilde{S}_n \leqslant x\}$. By the Fourier inversion formula,

$$\widetilde{F}_n(x) - \widetilde{\Phi}_3\Big(\frac{x-\mu}{\sigma}\Big) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,\frac{\widetilde{u}_n(t) - \widetilde{u}(t)}{-it}\,dt$$
$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,\frac{u_n(t) - u(t)}{-it}\,\omega(t)\,dt,$$

where the integrals are absolutely convergent. Hence

$$\widetilde{\Delta}_n \equiv \sup_x \Big|\widetilde{F}_n(x) - \widetilde{\Phi}_3\Big(\frac{x-\mu}{\sigma}\Big)\Big| \leqslant \frac{1}{2\pi} I, \tag{6.1}$$

where

$$I = \int_{-\infty}^{\infty}\Big|\frac{u_n(t) - u(t)}{t}\Big|\,|\omega(t)|\,dt.$$

Our purpose here is to properly estimate the last integral.

Near zero it can be estimated by virtue of Lemma 5.1 via (5.3). Indeed, changing the variable and using $|\omega(t)| \leqslant 1$, we have

$$I_0 \equiv \int_{|t|\leqslant 1/(\sigma L_3)}\Big|\frac{u_n(t) - u(t)}{t}\Big|\,|\omega(t)|\,dt \leqslant \int_{|t|\leqslant 1/L_3}\Big|\frac{f_n(t) - g(t)}{t}\Big|\,dt \leqslant cL_4. \tag{6.2}$$

For a similar integral I_1 over the complementary region, we have $I_1 \leqslant 2J + I_1'$ with

$$J = \int_{1/(\sigma L_3)}^{\infty}\Big|\frac{u_n(t)}{t}\Big|\,|\omega(t)|\,dt, \qquad I_1' = \int_{|t|>1/L_3}\Big|\frac{g(t)}{t}\Big|\,dt.$$

The last integral was estimated in Lemma 5.2. Using (6.2) and (5.6), we obtain

$$\widetilde{\Delta}_n \leqslant cL_4 + \frac{1}{\pi}\,J \tag{6.3}$$

with some absolute constant $c > 0$.

Let us now estimate J by involving the functional V. Putting $b^2 = \mathrm{V}(S_n)$, $b > 0$, we recall that $b^2 \leqslant \sigma^2$ and $a \equiv 1/(\sigma L_3) \leqslant 2$ as indicated in (4.3) and (5.5). We also note that

$$ab = \frac{b}{\sigma L_3} \geqslant \frac{b}{\sigma\sqrt{L_4}}. \tag{6.4}$$

Since the function $u_n(t)$ is (2π)-periodic, it makes sense to split the integration in the definition of J into the intervals

$$A_0 = (a, \pi), \quad A_k = \big((2k-1)\pi, (2k+1)\pi\big), \qquad k = 1, 2, \dots .$$

From (4.2) we have

$$|u_n(t)| \leqslant \exp\Big\{-2b^2\sin^2\Big(\frac{t}{2}\Big)\Big\}, \qquad t \in \mathbf{R}. \tag{6.5}$$

Next, $\sin(t/2) \geqslant \frac{t}{\pi}$ in $0 \leqslant t \leqslant \pi$, and so

$$|u_n(t)| \leqslant e^{-2b^2t^2/\pi^2}, \qquad t \in A_0.$$

Hence, putting $T = 2ab/\pi$, we find that

$$J_0 \equiv \int_{A_0} \left|\frac{u_n(t)}{t}\right| |\omega(t)|\, dt \leqslant \int_a^\infty e^{-2b^2t^2/\pi^2}\, \frac{dt}{t}$$

$$= \int_T^\infty e^{-s^2/2}\, \frac{ds}{s} < \frac{1}{T^2} \leqslant \frac{3}{(ab)^2} \leqslant \frac{3\sigma^2}{b^2} L_4, \tag{6.6}$$

where we have used (6.4).

The remaining integrals should be estimated in a different way by using the property that $\omega(t)$ is small when t is close to any integer multiple of 2π. Namely, by (6.5),

$$J_k \equiv \int_{A_k} \left|\frac{u_n(t)}{t}\right| |\omega(t)|\, dt \leqslant \frac{2}{((2k-1)\pi)^2} K,$$

where

$$K = \int_{-\pi}^{\pi} \exp\left\{-2b^2 \sin^2\left(\frac{t}{2}\right)\right\} \left|\sin \frac{t}{2}\right| dt$$

$$= 4 \int_0^{\pi/2} \exp\{-2b^2 \sin^2 s\} \sin s\, ds \leqslant 4 \int_0^{\pi/2} \exp\left\{-\frac{8b^2 s^2}{\pi^2}\right\} s\, ds < \frac{\pi^2}{4b^2}.$$

It follows that $\sum_{k=1}^\infty J_k < \frac{1}{b^2}$, and, together with (6.6), we have

$$J < \frac{3\sigma^2}{b^2} L_4 + \frac{1}{b^2} \leqslant \frac{7\sigma^2}{b^2} L_4,$$

where we use $\sigma^2 L_4 \geqslant \frac{1}{4}$; cf. (5.5). Combining this with (6.3), we have

$$\widetilde{\Delta}_n \leqslant \frac{c\sigma^2}{b^2} L_4. \tag{6.7}$$

It remains to apply Lemma 5.3, together with (5.4), (5.5), and (6.7), and use definition (6.1). Hence

$$\Delta_n \equiv \sup_x \left|\widetilde{F}_n(x) - \Phi_3\left(\frac{x-\mu}{\sigma}\right)\right| \leqslant \frac{c\sigma^2}{b^2} L_4 + cL_4(1 + L_4^{1/2})$$

with some absolute constant $c > 0$. If $L_4 \leqslant 1$, we get (4.5) in view of (5.7) and since $\sigma^2/b^2 \geqslant 1$. Otherwise, the required inequality

$$\Delta_n \leqslant \frac{c\sigma^2}{b^2} L_4$$

also holds with a sufficiently large c. Indeed, if $L_4 \geqslant 1$, then by (1.3),

$$|\Phi_3(x)| \leqslant 1 + |l_3| \leqslant 1 + L_3 \leqslant 1 + \sqrt{L_4} \leqslant 2L_4.$$

Thus, $\Delta_n \leqslant 3L_4$.

REFERENCES

[1] V. Becher, Y. Bugeaud, and T. A. Slaman, *The irrationality exponents of computable numbers*, Proc. Amer. Math. Soc., 144 (2016), pp. 1509–1521, https://doi.org/10.1090/proc/12841.

[2] S. N. Bernstein, *On a modification of Chebyshev's inequality and of the error formula of Laplace formula*, Uch. zap. n.-i., kafedr Ukraini, Otd. matem, no. 1, 1 (1924), pp. 38–49 (in Russian); in Collected Works: Vol. IV: Probability Theory: Mathematical Statistics (1911–1946), Nauka, Moscow, 1964, pp. 71–79 (in Russian).

[3] S. N. Bernstein, *Reconsideration of the question of the accuracy of the limit formula of Laplace*, Izv. Akad. Nauk SSSR Ser. Mat., 7 (1943), pp. 3–16; in Collected Works: Vol. IV: Probability Theory: Mathematical Statistics (1911–1946), Nauka, Moscow, 1964, pp. 396–408 (in Russian).

[4] A. Bikjalis, *Estimates of the remainder term in the central limit theorem*, Litovsk. Mat. Sb., 6 (1966), pp. 323–346 (in Russian).

[5] S. G. Bobkov, *Asymptotic expansions for products of characteristic functions under moment assumptions of non-integer orders*, in Convexity and Concentration, IMA Vol. Math. Appl. 161, Springer, New York, 2017, pp. 297–357, https://doi.org/10.1007/978-1-4939-7005-6_11.

[6] S. G. Bobkov, *Central limit theorem and Diophantine approximations*, J. Theoret. Probab., 31 (2018), pp. 2390–2411, https://doi.org/10.1007/s10959-017-0770-4.

[7] P. Tchebycheff, *Sur deux théorèmes relatifs aux probabilités*, Acta Math., 14 (1890), pp. 305–315, https://doi.org/10.1007/BF02413327.

[8] T. T. Cheng, *The normal approximation to the Poisson distribution and a proof of a conjecture of Ramanujan*, Bull. Amer. Math. Soc., 55 (1949), pp. 396–401, https://doi.org/10.1090/S0002-9904-1949-09223-6.

[9] H. Cramér, *On the composition of elementary errors* II: *Statistical applications*, Scand. Actuar. J., 11 (1928), pp. 141–180, https://doi.org/10.1080/03461238.1928.10416872.

[10] H. Cramér, *Random Variables and Probability Distributions*, 2nd ed., Cambridge Tracts in Math. and Math. Phys. 39, Cambridge Univ. Press, New York, 1962.

[11] P. Deheuvels, M. L. Puri, and S. S. Ralescu, *Asymptotic expansions for sums of nonidentically distributed Bernoulli random variables*, J. Multivariate Anal., 28 (1989), pp. 282–303, https://doi.org/10.1016/0047-259X(89)90111-5.

[12] C.-G. Esseen, *Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law*, Acta Math., 77 (1945), pp. 1–125, https://doi.org/10.1007/BF02392223.

[13] W. Feller, *On the normal approximation to the binomial distribution*, Ann. Math. Statistics, 16 (1945), pp. 319–329, https://doi.org/10.1214/aoms/1177731058.

[14] F. Götze and A. Yu. Zaitsev, *Explicit rates of approximation in the CLT for quadratic forms*, Ann. Probab., 42 (2014), pp. 354–397, https://doi.org/10.1214/13-AOP839.

[15] B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, 2nd ed., Addison-Wesley, Reading, MA, 1968.

[16] V. G. Mikhailov, *On a refinement of the central limit theorem for sums of independent random indicators*, Theory Probab. Appl., 38 (1993), pp. 479–489, https://doi.org/10.1137/1138044.

[17] L. V. Osipov, *Asymptotic expansions of the distribution function of a sum of independent lattice random variables*, Theory Probab. Appl., 14 (1969), pp. 450–457, https://doi.org/10.1137/1114057.

[18] V. V. Petrov, *Sums of Independent Random Variables*, Ergeb. Math. Grenzgeb. 82, Springer-Verlag, New York, 1975, https://doi.org/10.1007/978-3-642-65809-9.

[19] V. V. Petrov, *Limit Theorems of Probability Theory. Sequences of Independent Random Variables*, Oxford Stud. Probab. 4, The Clarendon Press, New York, 1995.

[20] Yu. V. Prokhorov and V. V. Ulyanov, *Some approximation problems in statistics and probability*, in Limit Theorems in Probability, Statistics and Number Theory, Springer Proc. Math. Stat. 42, Springer, Heidelberg, 2013, pp. 235–249, https://doi.org/10.1007/978-3-642-36068-8_11.

[21] V. Statuljavičjus, *On the asymptotic expansion of the characteristic function of a sum of independent random variables*, Litovsk. Mat. Sb., 2 (1962), pp. 227–232 (in Russian).

[22] V. A. Statuljavičus, *Limit theorems for densities and asymptotic expansions for distributions of sums of independent random variables*, Theory Probab. Appl., 10 (1965), pp. 582–595, https://doi.org/10.1137/1110074.

[23] V. V. Ulyanov, *Asymptotic expansions for distributions of sums of independent random variables in H*, Theory Probab. Appl., 31 (1987), pp. 25–39, https://doi.org/10.1137/1131002.

[24] J. V. Uspensky, *Introduction to Mathematical Probability*, McGraw-Hill, New York, 1937.

[25] A. Yu. Volkova, *A refinement of the central limit theorem for sums of independent random indicators*, Theory Probab. Appl., 40 (1995), pp. 791–794, https://doi.org/10.1137/1140093.