Discussiones Mathematicae
Probability and Statistics 31 (2011) 29–39
doi:10.7151/dmps.1138

ON THE UNIVERSAL CONSTANT IN THE KATZ-PETROV AND OSIPOV INEQUALITIES *

Victor Korolev

Faculty of Computational Mathematics and Cybernetics
Moscow State University
Institute for Informatics Problems
Russian Academy of Sciences

e-mail: vkorolev@cmc.msu.su

and

Sergey Popov

Faculty of Computational Mathematics and Cybernetics
Moscow State University

Abstract

Upper estimates are presented for the universal constant in the Katz-Petrov and Osipov inequalities which do not exceed 3.1905.

Keywords and phrases: central limit theorem, convergence rate estimate, absolute constant, Katz-Petrov inequality, Osipov inequality.

2010 Mathematics Subject Classification: 60F05, 60G50.

*Research supported by Russian Foundation for Basic Research, projects 11-01-00515 and 11-07-00112 and the Ministry for Science and Education of Russian Federation, state contracts P958, P779 and 16.740.11.0133.

1. Introduction

Let $X_1, X_2, \ldots$ be independent random variables with $\mathsf{E}X_i = 0$ and $0 < \mathsf{E}X_i^2 \equiv \sigma_i^2 < \infty$, $i = 1, 2, \ldots$ For $n \in \mathbb{N}$ denote $S_n = X_1 + \cdots + X_n$, $B_n^2 = \sigma_1^2 + \cdots + \sigma_n^2$. Let $\Phi(x)$ be the standard normal distribution function,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-z^2/2} dz, \qquad x \in \mathbb{R}.$$

Denote

$$\Delta_n = \sup_y |\mathsf{P}(S_n < yB_n) - \Phi(y)|.$$

Let $\mathcal{G}$ be the class of real-valued functions $g(x)$ of $x \in \mathbb{R}$ such that

- $g(x)$ is even;
- $g(x)$ is non-negative for all x and $g(x) > 0$ for $x > 0$;
- $g(x)$ does not decrease for $x > 0$;
- the function $x/g(x)$ does not decrease for $x > 0$.

In 1963 M. Katz [4] proved that, whatever $g \in \mathcal{G}$ is, if the random variables $X_1, X_2, \ldots$ are identically distributed with $\mathsf{E}X_1^2 g(X_1) < \infty$, then there exists a finite positive absolute constant C such that

$$\Delta_n \leqslant C \cdot \frac{\mathsf{E}X_1^2 g(X_1)}{\sigma_1^2 g\big(\sigma_1\sqrt{n}\big)}. \tag{1}$$

In 1965 this result was generalized by V.V. Petrov [11] to the case of not necessarily identically distributed random variables (also see [12]): whatever $g \in \mathcal{G}$ is, if $\mathsf{E}X_i^2 g(X_i) < \infty$, $i = 1, \ldots, n$, then there exists a finite positive absolute constant C such that

$$\Delta_n \leqslant \frac{C}{B_n^2 g(B_n)} \sum_{i=1}^{n} \mathsf{E}X_i^2 g(X_i). \tag{2}$$

The present paper aims at giving an upper bound of the absolute constant C in (2). It will be shown that this bound does not depend on the particular form of $g \in \mathcal{G}$ (and, hence, is universal) and does not exceed 3.1905 in the general case. We also give sharper bounds for some special cases.

In particular, the function

$$g(x) = \min\{|x|,\, B_n\}, \quad x \in \mathbb{R},$$

is obviously in $\mathcal{G}$. In this case inequality (2) turns into

$$(2') \qquad \Delta_n \leqslant C\biggl(\frac{1}{B_n^2}\sum_{i=1}^{n} \mathsf{E}X_i^2\mathbb{I}(|X_i| \geqslant B_n) + \frac{1}{B_n^3}\sum_{i=1}^{n}\mathsf{E}|X_i|^3\mathbb{I}(|X_i| < B_n)\biggr).$$

This inequality was proved in 1966 by L.V. Osipov [7] (also see [12], Ch. V, Section 3, Theorem 8). In [8, 9] L. Paditz showed that in (2′) $C < 4.77$. In 1986 he also noted [10] that with the account of Lemma 12.2 in [1] the techniques used in [8, 9] makes it possible to lower this estimate down to $C < 3.51$. Apparently, being unaware of the result of Paditz, in 2001 Chen and Shao published the paper [2] in which by the Tikhomirov-Stein method inequality (2′) was re-proved with $C = 4.1$.

From the results of the present paper it follows that the estimates of the constant C in (2′) can be sharpened to at least $C \leqslant 3.1905$.

2. Auxiliary statements

Lemma 1. *Let X be a random variable with $\mathsf{E}|X|^3 < \infty$ and $\mathsf{E}X = a$. Let*

$$K = \frac{17 + 7\sqrt{7}}{27} \approx 1.315565\ldots$$

Then

$$\mathsf{E}|X - a|^3 \leqslant \min\bigl\{K\mathsf{E}|X|^3,\ \mathsf{E}|X|^3 + 3|a|\mathsf{E}X^2 + a^2\mathsf{E}|X|\bigr\}.$$

Proof. On the one hand, it is obvious that

$$\begin{aligned}\mathsf{E}|X - a|^3 &= \mathsf{E}|X - a|(X - a)^2 = \mathsf{E}|X - a|(X^2 - 2aX + a^2) \leqslant \\ &= \mathsf{E}|X|^3 - 2a\mathsf{E}(X|X|) + a^2\mathsf{E}|X| + |a|\mathsf{E}X^2 - 2|a|a\mathsf{E}X + |a|^3 \leqslant \\ &= \mathsf{E}|X|^3 + 3|a|\mathsf{E}X^2 + a^2\mathsf{E}|X|.\end{aligned}$$

On the other hand, using the result of [3] stating that the extremum of a functional linear in the distribution function of the random variable X under the single linear moment-type condition $\mathsf{E}X = a$ is attained at some two-point distribution, in [6] (see Lemma 5 there) it was proved that

$$\sup_{X\colon \mathsf{E}|X|^3<\infty} \frac{\mathsf{E}|X-\mathsf{E}X|^3}{\mathsf{E}|X|^3} = \frac{17+7\sqrt{7}}{27} < 1.3156,$$

that completes the proof.

Lemma 2. 1°. *Let $q>0$. Then*

$$\sup_x |\Phi(qx)-\Phi(x)| \leqslant \frac{1}{\sqrt{2\pi e}}\Big(\max\Big\{q,\,\frac{1}{q}\Big\}-1\Big).$$

2°. *Let $a\in\mathbb{R}$. Then*

$$\sup_x |\Phi(x+a)-\Phi(x)| \leqslant \frac{|a|}{\sqrt{2\pi}}.$$

The simple proof of this lemma is based on the Lagrange formula (also see [12], Chapter 5).

Lemma 3. *Let X be a random variable with $\mathsf{E}X=0$ and $\mathsf{E}X^2=1$. Then*

$$\sup_x |\mathsf{P}(X<x)-\Phi(x)| \leqslant 0.541.$$

For the proof see, e.g., Lemma 12.2 in [1].

3. Main result

Theorem.

1°. *Let $g\in\mathcal{G}$, $n\geqslant 1$ be an integer, random variables $X_1,\ldots,X_n$ be independent with $\mathsf{E}X_i=0$ and $\mathsf{E}X_i^2 g(X_i)<\infty$, $i=1,\ldots,n$. Then inequality* (2) *holds with $C\leqslant 3.1905$.*

2°. *Let, in addition to the conditions specified in* 1°, *the random variables $X_1,\ldots,X_n$ be identically distributed. Then inequality* (1) *holds with $C\leqslant 3.0466$.*

3°. *Let, in addition to the conditions specified in* 1°, *the random variables* $X_1, \dots, X_n$ *have symmetric distributions. Then inequality* (2) *holds with* $C \leqslant 2.0409$.

4°. *Let, in addition to the conditions specified in* 2°, *the random variables* $X_1, \dots, X_n$ *have symmetric distribution. Then inequality* (1) *holds with* $C \leqslant 1.9363$.

Proof. Following the mainstream of the proof of (2) in [12], we will slightly adjust it to our purposes.

1°. Consider the truncated random variables

$$\widetilde{X}_j = X_j \mathbb{I}(|X_j| < B_n), \quad j = 1, 2, \dots,$$

where $\mathbb{I}(A)$ is the indicator function of an event A: if ω is an elementary outcome, then

$$\mathbb{I}(A) = \mathbb{I}(\omega, A) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

For integer $j \geqslant 1$ and $n \geqslant 1$ denote

$$\widetilde{a}_j = \mathsf{E}\widetilde{X}_j, \quad \widetilde{A}_n = \widetilde{a}_1 + \dots + \widetilde{a}_n, \quad \widetilde{\sigma}_j^2 = \mathsf{D}\widetilde{X}_j,$$

$$\widetilde{B}_n^2 = \widetilde{\sigma}_1^2 + \dots + \widetilde{\sigma}_n^2, \quad F_j(x) = \mathsf{P}(X_j < x).$$

Since $\mathsf{E}X_j = 0$, then

$$\Bigg| \int\limits_{|x|<B_n} x dF_j(x) \Bigg| = \Bigg| \int\limits_{|x|\geqslant B_n} x dF_j(x) \Bigg|. \tag{3}$$

Let $\alpha \in (0, 1)$. Assume that $\widetilde{B}_n^2 \leqslant \alpha B_n^2$. Then with the account of (3) we have

$$(1-\alpha)B_n^2 \leqslant B_n^2 - \widetilde{B}_n^2 = \sum_{j=1}^n \int\limits_{|x|<B_n} x^2 dF_j(x) + \sum_{j=1}^n \int\limits_{|x|\geqslant B_n} x^2 dF_j(x)$$

$$-\sum_{j=1}^{n}\int\limits_{|x|<B_n} x^2 dF_j(x)+\sum_{j=1}^{n}\Big(\int\limits_{|x|<B_n} x dF_j(x)\Big)^2$$

$$=\sum_{j=1}^{n}\int\limits_{|x|\geqslant B_n} x^2 dF_j(x)+\sum_{j=1}^{n}\Big(\int\limits_{|x|\geqslant B_n} x dF_j(x)\Big)^2\leqslant 2\sum_{j=1}^{n}\int\limits_{|x|\geqslant B_n} x^2 dF_j(x) \tag{4}$$

$$=2\sum_{j=1}^{n}\int\limits_{|x|\geqslant B_n}\frac{x^2 g(x)}{g(x)}dF_j(x)\leqslant\frac{2}{g(B_n)}\sum_{j=1}^{n}\mathsf{E}X_j^2 g(X_j).$$

This means that, if $\widetilde{B}_n^2\leqslant\alpha B_n^2$, then

$$\frac{1}{B_n^2 g(B_n)}\sum_{j=1}^{n}\mathsf{E}X_j^2 g(X_j)\geqslant\frac{1-\alpha}{2}. \tag{5}$$

From now on we will assume that

$$\widetilde{B}_n^2>\alpha B_n^2. \tag{6}$$

Denote $Y_n=\widetilde{X}_1+\dots+\widetilde{X}_n$. The event $\{S_n<xB_n\}$ implies the event

$$\{Y_n<xB_n\}\bigcup\{|X_1|\geqslant B_n\}\bigcup\cdots\bigcup\{|X_n|\geqslant B_n\},$$

whereas the event $\{Y_n<xB_n\}$ implies the event

$$\{S_n<xB_n\}\bigcup\{|X_1|\geqslant B_n\}\bigcup\cdots\bigcup\{|X_n|\geqslant B_n\}.$$

Therefore

$$\Delta_n\leqslant Q_1+Q_2+Q_3,$$

where

$$Q_1 = \sup_x \left| \mathsf{P}\left(\frac{Y_n - \widetilde{A}_n}{\widetilde{B}_n} < \frac{xB_n - \widetilde{A}_n}{\widetilde{B}_n} \right) - \Phi\left(\frac{xB_n - \widetilde{A}_n}{\widetilde{B}_n} \right) \right|,$$

$$Q_2 = \sup_x \left| \Phi\left(\frac{xB_n - \widetilde{A}_n}{\widetilde{B}_n} \right) - \Phi(x) \right|, \qquad Q_3 = \sum_{j=1}^{n} \mathsf{P}(|X_j| \geqslant B_n).$$

By virtue of the Berry-Esseen inequality with the best known upper bound of the absolute constant [13] with the account of Lemma 1 and condition (6) we have

$$\begin{aligned} Q_1 &\leqslant \frac{0.56}{\widetilde{B}_n^3} \sum_{j=1}^{n} \mathsf{E}|\widetilde{X}_j - \widetilde{a}_j|^3 \leqslant \frac{0.56 \cdot 1.3156}{\alpha^{3/2} B_n^3} \sum_{j=1}^{n} \mathsf{E}|\widetilde{X}_j|^3 \\ &\leqslant \frac{0.736736}{\alpha^{3/2} B_n^3} \sum_{j=1}^{n} \int\limits_{|x|<B_n} \frac{|x|}{g(x)} x^2 g(x) dF_j(x) \qquad (7) \\ &\leqslant \frac{0.736736}{\alpha^{3/2} B_n^2 g(B_n)} \sum_{j=1}^{n} \mathsf{E} X_j^2 g(X_j). \end{aligned}$$

We obviously have

$$Q_2 \leqslant Q_{21} + Q_{22},$$

where

$$\begin{aligned} Q_{21} &= \sup_x \left| \Phi\left(xB_n/\widetilde{B}_n\right) - \Phi(x) \right|, \\ Q_{22} &= \sup_x \left| \Phi\left(x - \widetilde{A}_n/\widetilde{B}_n\right) - \Phi(x) \right|. \end{aligned}$$

Furthermore, by virtue of Lemma 2 (1°) and condition (6) we obtain

$$Q_{21} \leqslant \frac{1}{\sqrt{2\pi e}} \left(\frac{B_n}{\widetilde{B}_n} - 1 \right) = \frac{B_n^2 - \widetilde{B}_n^2}{\sqrt{2\pi e} \widetilde{B}_n (B_n + \widetilde{B}_n)} \leqslant \frac{B_n^2 - \widetilde{B}_n^2}{\sqrt{2\pi e \alpha} (1 + \sqrt{\alpha}) B_n^2}.$$

Estimating the difference $B_n^2 - \widetilde{B}_n^2$ in the numerator in the same way as we did to establish relation (4), we appear at the inequality

$$Q_{21} \leqslant \frac{2}{\sqrt{2\pi e\alpha}(1+\sqrt{\alpha})B_n^2 g(B_n)} \sum_{j=1}^{n} \mathsf{E}X_j^2 g(X_j). \tag{8}$$

By virtue of Lemma 2 (2°) and conditions (6) and (3) we obtain

$$\begin{aligned} Q_{22} &\leqslant \frac{|\widetilde{A}_n|}{\sqrt{2\pi}\widetilde{B}_n} \\ &\leqslant \frac{1}{\sqrt{2\pi\alpha}B_n} \sum_{j=1}^{n} \Bigg| \int\limits_{|x|<B_n} x dF_j(x) \Bigg| = \frac{1}{\sqrt{2\pi\alpha}B_n} \sum_{j=1}^{n} \Bigg| \int\limits_{|x|\geqslant B_n} x dF_j(x) \Bigg| \qquad (9) \\ &\leqslant \frac{1}{\sqrt{2\pi\alpha}B_n} \sum_{j=1}^{n} \int\limits_{|x|\geqslant B_n} \frac{x^2 g(x)}{|x|g(x)} dF_j(x) \leqslant \frac{1}{\sqrt{2\pi\alpha}B_n^2 g(B_n)} \sum_{j=1}^{n} \mathsf{E}X_j^2 g(X_j). \end{aligned}$$

Unifying (8) and (9) we obtain

$$Q_2 \leqslant \frac{1}{\sqrt{2\pi\alpha}} \left(1 + \frac{2}{\sqrt{e}(1+\sqrt{\alpha})}\right) \cdot \frac{1}{B_n^2 g(B_n)} \sum_{j=1}^{n} \mathsf{E}X_j^2 g(X_j). \tag{10}$$

Finally, by the Markov inequality we have

$$Q_3 \leqslant \frac{1}{B_n^2 g(B_n)} \sum_{j=1}^{n} \mathsf{E}X_j^2 g(X_j). \tag{11}$$

From (7), (10) and (11) it follows that, under condition (6),

$$\Delta_n \leqslant \frac{C_1(\alpha)}{B_n^2 g(B_n)} \sum_{j=1}^{n} \mathsf{E}X_j^2 g(X_j) \tag{12}$$

with

$$C_1(\alpha) = \frac{0.736736}{\alpha^{3/2}} + \frac{1}{\sqrt{2\pi\alpha}} \left(1 + \frac{2}{\sqrt{e}(1+\sqrt{\alpha})}\right) + 1. \tag{13}$$

To choose the optimal value of α and, hence, $C_1(\alpha)$ note that $C_1(\alpha)$ is a decreasing function of $\alpha \in (0,1)$. On the other hand, for the inequality (12) to be reasonable irrespective of condition (6), that is, for all possible distributions of X_j, the parameter α should be chosen so that for distributions with $\widetilde{B}_n^2 \leqslant \alpha B_n^2$ estimate (12) becomes trivial. Thus, with the account of Lemma 3 and relation (5) we arrive at the conclusion that the optimal α and $C_1(\alpha)$ must be tied up by the equation

$$C_1(\alpha) = \frac{2 \cdot 0.541}{1-\alpha}. \tag{14}$$

The left-hand side of this equation is decreasing in α whereas its right-hand side increases. Therefore, equation (14) has the unique solution $\alpha_1 \approx 0.66086$ providing $C_1(\alpha_1) \approx 3.19045\ldots$ Item 1° is thus proved.

2°. The proof of this statement is a word-for-word copy of the proof of 1° with the only change: the coefficient 0.56 in (7) should be replaced by the coefficient 0.4784 which is the best known upper bound of the constant in the Berry-Esseen inequality for sums of independent identically distributed random variables [5]. So, instead of (14), the equation

$$C_2(\alpha) = \frac{2 \cdot 0.541}{1-\alpha}. \tag{15}$$

should be solved with

$$C_2(\alpha) = \frac{0.62938304}{\alpha^{3/2}} + \frac{1}{\sqrt{2\pi\alpha}}\left(1 + \frac{2}{\sqrt{e}(1+\sqrt{\alpha})}\right) + 1 \tag{16}$$

yielding the solution $\alpha_2 \approx 0.64484$ and $C_2(\alpha_2) \approx 3.046506\ldots$

3°. In this case the expectations of the summands equal zero. Therefore, the coefficient 2 in (4) and, hence, in (8) as well as the coefficient 1.3156 in (7) turn into 1 whereas Q_{22} vanishes. Therefore, the optimal value of α should be sought as the solution to the equation

$$C_3(\alpha) = \frac{0.541}{1-\alpha}, \tag{17}$$

where

$$C_3(\alpha) = \frac{0.56}{\alpha^{3/2}} + \frac{1}{\sqrt{2\pi e\alpha}(1+\sqrt{\alpha})} + 1. \tag{18}$$

The unique solution of (17) is $\alpha_3 \approx 0.73491$ yielding $C_3(\alpha_3) \approx 2.04083\ldots$

4°. In this case the proof repeats the proof of 3° with $C_3(\alpha)$ replaced by

$$C_4(\alpha) = \frac{0.4784}{\alpha^{3/2}} + \frac{1}{\sqrt{2\pi e\alpha}(1+\sqrt{\alpha})} + 1. \tag{19}$$

The unique solution of the equation

$$C_4(\alpha) = \frac{0.541}{1-\alpha} \tag{20}$$

is $\alpha_4 \approx 0.720595$ providing $C_4(\alpha_4) \approx 1.93625\ldots$ The theorem is proved.

REFERENCES

[1] R.N. Bhattacharya and R. Ranga Rao, Normal Approximation and Asymptotic Expansions (New York, Wiley, 1976).

[2] L.H.Y. Chen and Q.M. Shao, *A non-uniform Berry-Esseen bound via Stein's method*, Probability Theory and Related Fields **120** (2001) 236–254.

[3] W. Hoeffding, *The extrema of the expected value of a function of independent random variables*, Ann. Math. Statist. **19** (1948) 239–325.

[4] M. Katz, *Note on the Berry-Esseen theorem*, Annals of Math. Statist. **39** (4) (1963) 1348–1349.

[5] V.Yu. Korolev and I.G. Shevtsova, *An improvement of the Berry-Esseen inequality with applications to Poisson and mixed Poisson random sums*, Scandinavian Actuarial Journal, 2010. Online first: http://www.informaworld.com/10.1080/03461238.2010.485370, 04 June 2010.

[6] J.S. Nefedova and I.G. Shevtsova, *On non-uniform estimates of convergence rate in the central limit theorem*, Theory Probab. Appl. **56** (2011), to appear.

[7] L.V. Osipov, *A refinement of the Lindeberg theorem*, Theory Probab. Appl. **11** (2) (1966) 339–342.

[8] L. Paditz, *Bemerkungen zu einer Fehlerabschätzung im zentralen Grenzwertsatz*, Wiss. Z. Hochschule für Verkehrswesen Friedrich List **27** (4) (1980) 829–837.

[9] L. Paditz, *On error-estimates in the central limit theorem for generalized linear discounting*, Math. Operationsforsch. u. Statist., Ser. Statistics **15** (4) (1984) 601–610.

[10] L. Paditz, *Über eine Fehlerabschätzung im zentralen Grenzwertsatz*, Wiss. Z. Hochschule für Verkehrswesen Friedrich List Dresden **33** (2) (1986) 399–404.

[11] V.V. Petrov, *An estimate of the deviation of the distribution of a sum of independent random variables from the normal law*, Soviet Math. Dokl. **160** (5) (1965) 1013–1015.

[12] V.V. Petrov, Sums of Independent Random Variables (New York, Springer, 1975).

[13] I.G. Shevtsova, *A refinement of the estimates of the rate of convergence in the Lyapunov theorem*, Doklady Mathematics **435** (1) (2010) 26–28.

Received 16 February 2011