Liet. mat. rink. LMD darbai, **48/49**, 2008, 389–394

Error estimate in the central limit theorem

Aurelija KASPARAVIČIŪTĖ, Kazimieras PADVELSKIS (VDU)
e-mail: aurelija2@one.lt, k.padvelskis@if.vdu.lt

Abstract. In this paper, we determined, independent identically distributed random variable's $\{X_k, k = 1, 2, \ldots\}$ centered and normalized sum's $S_n = \sum_{k=1}^{n} X_k$ distribution's $F_n(x) = \mathbf{P}(Z_n < x)$ exact error estimate in case of the normal approximation with one Čebyšova's asymptotic expansion's term.

Keywords: the Central limit theorem, normal distribution, error estimate, cumulants.

1. Introduction

Let $\{X_k, k = 1, 2, \ldots\}$ – independent identically distributed random variables (r.v.), having absolute finite moments

$$\beta_s = \mathbf{E}\left|X_k\right|^s < \infty, \quad s = 1, 2, \ldots, \ k = 1, 2, \ldots, n. \tag{1}$$

Note, $\mathbf{E}X_1 = 0$, $\mathbf{D}X_1 = \sigma^2$ are mean and variance. Besides, it is assumed, that

$$\left|\mathbf{E}X_1^s\right| \leqslant s!K^{s-3}\beta_3, \quad K > 0, \ s \geqslant 4. \tag{2}$$

Note $S_n = \sum_{k=1}^{n} X_k$ – r.v. sum, then $\mathbf{E}S_n = 0$ and $\mathbf{D}S_n = n\sigma^2$. Let

$$Z_n = \left(\sigma\sqrt{n}\right)^{-1} S_n, \quad F_n(x) = \mathbf{P}\left(Z_n < x\right), \tag{3}$$

be sum's Z_n distribution function that Fourier transform $f_n(t) = \mathbf{E}\exp\{itZ_n\}$.

The aim of this paper is to determine distribution's (3) error estimate in case of the normal approximation with one Čebyšova's asymptotic expansion's (see [3]) term that depends upon the third r.v. cumulant $\gamma_3 = \Gamma_3(X_1)$, scilicet

$$G_n(x) = \Phi(x) + \frac{\mathrm{e}^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \frac{\gamma_3}{6\sigma^3\sqrt{n}}\left(1 - x^2\right), \tag{4}$$

here $\Phi(x)$ – the standard normal distribution function.

We considered two cases, then r.v. are both absolutely continuous, both lattice – their values achievable with positive probabilities belong to arithmetical progression $\{a + lh, l = 0, \pm 1, \ldots\}$, here $a \in \mathbf{R}$ and $h > 0$ – maximal arithmetical progression's difference.

If independent identically distributed r.v. are lattice, then function (4) is supplemented with such periodic function $\mathrm{S}(x) = [x] - x + 1/2$(see [4], [5]), here $[x]$ –

integer part of a number, scilicet

$$G_n^*(x) = \Phi(x) + \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi n}} \frac{\gamma_3}{6\sigma^3}\big(1 - x^2\big) + \frac{h e^{-\frac{1}{2}x^2}}{\sigma\sqrt{2\pi n}} S\big((\sigma x\sqrt{n} - an)h^{-1}\big). \tag{5}$$

It should be remarked on, that such type of value of an error estimate in the Central limit theorems has been made in such books, like [3]–[5], but in all of them have been proofed, that the speed of convergence is $o(1/\sqrt{n})$ order.

We will present you some new results, wherein absolute constants are evaluated too.

Moreover, in this paper the method of characteristic's functions is used to find the remainder term of the normal approximation with one Čebyšova's asymptotic expansion's term.

Therefore at first we specified smoothing inequalities.

2. Error estimate of the normal approximation

THEOREM 2.1. 1. *If $F(x)$ – r.v. X_1 distribution function, and function $G_1(x)$ is described by* (4) *formula, besides these function's Fourier transforms are $f(t)$ and $g(t)$, then for every $T > 0$*

$$\sup_x \big|F(x) - G_1(x)\big| \leqslant \frac{3}{10}\int_{-T}^{T}\left|\frac{f(t) - g(t)}{t}\right| \mathrm{d}t + 2\frac{\beta_3}{T\sigma^3}. \tag{6}$$

2. *Assume that X_1 is lattice r.v. If $F(x)$ – lattice r.v. X_1 distribution function, and function $G_1^*(x)$ is described by* (5) *formula, besides these function's Fourier transforms are $f(t)$ and $g^*(t)$, then for every $T > 0$*

$$\sup_x \big|F(x) - G_1^*(x)\big| \leqslant \frac{3}{10}\int_{-T}^{T}\left|\frac{f(t) - g^*(t)}{t}\right| \mathrm{d}t + \frac{(1 \vee h)\beta_3}{T\sigma^3}, \tag{7}$$

here $1 \vee h = \max\{1, h\}$.

This theorem was proofed using the proof placed in [4].

It's easy to obtain (4), (5) functions Fourier transforms, respectively $g(t)$, $g^*(t)$.

If r.v. are absolutely continuous, then

$$g_n(t) = e^{-\frac{1}{2}t^2} + \frac{\gamma_3 (it)^3}{6\sigma^3\sqrt{n}} e^{-\frac{1}{2}t^2}, \tag{8}$$

and in case of lattice r.v., we obtained that

$$g_n^*(t) = e^{-\frac{1}{2}t^2} + \frac{\gamma_3 (it)^3}{6\sigma^3\sqrt{n}} e^{-\frac{1}{2}t^2} - \frac{ht}{2\pi\sigma\sqrt{n}} {\sum_{v=-\infty}^{\infty}}' \frac{e^{-ih^{-1}2\pi v a n}}{v} \exp\Big\{-\frac{1}{2}\big(t + h^{-1}2\pi v\sigma\sqrt{n}\big)^2\Big\}. \tag{9}$$

Assume that considered r.v. $\{X_k, k = 1, 2, \ldots\}$ are absolutely continuous.

THEOREM 2.2. *If independent identically distributed r.v.* $\{X_k, k = 1, 2, \dots\}$ *have* (1) *absolute finite moments, and the* (2) *condition is satisfied, then*

$$\sup_x \bigl|F_n(x) - G_n(x)\bigr| \leqslant 6\bigl((K+2) \vee \sqrt{2\beta_3}\bigr)^2 \frac{\beta_3^2}{n\sigma^6}, \tag{10}$$

here $(K+2) \vee \sqrt{2\beta_3} = \max\{K+2, \sqrt{2\beta_3}\}$.

Proof. At first we have to evaluate $|f_n(t) - g_n(t)|$. According to the inequality $|e^x - 1 - y| \leqslant |x - y| \exp\{|x-y| + |y|\} + (1/2)|y|^2 e^{|y|}$, we have

$$\begin{aligned}
\bigl|f_n(t) - g_n(t)\bigr| &= e^{-\frac{1}{2}t^2}\left|\exp\left\{\ln f_n(t) + \frac{1}{2}t^2\right\} - 1 - \frac{\gamma_3(it)^3}{6\sigma^3\sqrt{n}}\right| \\
&\leqslant \left|\ln f_n(t) + \frac{1}{2}t^2 - \frac{\gamma_3(it)^3}{6\sigma^3\sqrt{n}}\right| \\
&\quad \times \exp\left\{-\frac{1}{2}t^2 + \left|\ln f_n(t) + \frac{1}{2}t^2 - \frac{\gamma_3(it)^3}{6\sigma^3\sqrt{n}}\right| + \left|\frac{\gamma_3(it)^3}{6\sigma^3\sqrt{n}}\right|\right\} \\
&\quad + \frac{1}{2}\left|\frac{\gamma_3(it)^3}{6\sigma^3\sqrt{n}}\right|^2 \exp\left\{-\frac{1}{2}t^2 + \left|\frac{\gamma_3(it)^3}{6\sigma^3\sqrt{n}}\right|\right\}.
\end{aligned}$$

Well known that $f_n(t) = f_{X_1}^n(t(\sigma\sqrt{n})^{-1})$, then

$$\ln f_n(t) = n\sum_{s=2}^{\infty} \frac{\gamma_s}{s!}\bigl(it(\sigma\sqrt{n})^{-1}\bigr)^s. \tag{11}$$

Besides, if the (2) condition is satisfied, then it is easy to show (see [1], [2]) that

$$|\gamma_s| \leqslant s!(3/2)\bigl((K+2) \vee \sqrt{2\beta_3}\bigr)^{s-2}\beta_3, \quad s \geqslant 4,\ K > 0. \tag{12}$$

Well known that $\Gamma_s(S_n) = \sum_{k=1}^n \Gamma_s(X_k) = n\gamma_s$, here $\gamma_s = \Gamma_s(X_1)$, and

$$\Gamma_s\bigl(S_{n(\sigma\sqrt{n})^{-1}}\bigr) = \frac{\Gamma_s(S_n)}{\sigma^s(\sqrt{n})^s} = \frac{\gamma_s}{\sigma^s(\sqrt{n})^{s-2}}. \tag{13}$$

Consequently, if the (12) condition is satisfied, then

$$\bigl|\Gamma_s(Z_n)\bigr| \leqslant s!(3/2)\bigl((K+2) \vee \sqrt{2\beta_3}\bigr)^{s-2}\beta_3\bigl(\sigma^s(\sqrt{n})^{s-2}\bigr)^{-1}. \tag{14}$$

Reference to (11), (12), (14) it's easy to evaluate

$$\begin{aligned}
&\Bigl|\ln f_n(t) + (1/2)t^2 - \gamma_3(it)^3\bigl(6\sigma^3\sqrt{n}\bigr)^{-1}\Bigr| \\
&\leqslant (3/2)\bigl[(K+2) \vee \sqrt{2\beta_3}\bigr]^2\beta_3|t|^4(n\sigma^4)^{-1}\sum_{s=4}^{\infty}\Bigl[|t|\bigl((K+2) \vee \sqrt{2\beta_3}\bigr)(\sigma\sqrt{n})^{-1}\Bigr]^{s-4}
\end{aligned}$$

$$\leqslant (15/2)\big[(K+2)\vee\sqrt{2\beta_3}\big]^2\beta_3^2|t|^4(\sigma^6 n)^{-1};$$

$$\big|\gamma_3(it)^3(6\sigma^3\sqrt{n})^{-1}\big| \leqslant (1/4)\big[(K+2)\vee\sqrt{2\beta_3}\big]\beta_3|t|^3(\sigma^6 n)^{-1},$$

if $|t|\leqslant 4\sigma^3\sqrt{n}[5((K+2)\vee\sqrt{2\beta_3})\beta_3]^{-1}$ and $\beta_3\geqslant 1$. Therefore

$$\big|f_n(t)-g_n(t)\big| \leqslant \big((K+2)\vee\sqrt{2\beta_3}\big)^2\beta_3^2(\sigma^6 n)^{-1} \times\Big[(15/2)|t|^4\mathrm{e}^{-\frac{9}{2}t^2}+(1/16)|t|^6\mathrm{e}^{-\frac{3}{10}t^2}\Big].$$

Now we have to use (6) smoothing inequality than $T=4\sigma^3\sqrt{n}[5((K+2)\vee\sqrt{2\beta_3})^2\beta_3]^{-1}$ and $T_1=4\sigma^3 n[5((K+2)\vee\sqrt{2\beta_3})^2\beta_3]^{-1}$, scilicet

$$\sup_x\big|F_n(x)-G_n(x)\big| \leqslant \frac{3}{10}(\mathrm{I}_1+\mathrm{I}_2)+2\beta_3(\sigma^3 T)^{-1}, \tag{15}$$

here

$$I_1=\int_{|t|\leqslant T}\Big|t^{-1}\big(f_n(t)-g_n(t)\big)\Big|\,\mathrm{d}t \leqslant I_{1.1}+I_{1.2}, \tag{16}$$

and

$$I_2=\int_{T<|t|\leqslant T_1}\Big|t^{-1}\big(f_n(t)-g_n(t)\big)\Big|\,\mathrm{d}t \leqslant I_{2.1}+I_{2.2}. \tag{17}$$

At first, evaluate I_1 integral. Considering to

$$\begin{aligned}
I_{1.1} &= (15/2)\big[(K+2)\vee\sqrt{2\beta_3}\big]^2\beta_3^2(\sigma^6 n)^{-1}\int_{-T}^{T}|t|^3\mathrm{e}^{-\frac{9}{2}t^2}\,\mathrm{d}t\\
&\leqslant (10/27)\big[(K+2)\vee\sqrt{2\beta_3}\big]^2\beta_3^2(\sigma^6 n)^{-1}\int_0^{\infty}u\mathrm{e}^{-u}\,\mathrm{d}u\\
&= (10/27)\big[(K+2)\vee\sqrt{2\beta_3}\big]^2\beta_3^2(\sigma^6 n)^{-1};\\
I_{1.2} &= (1/16)\big[(K+2)\vee\sqrt{2\beta_3}\big]^2\beta_3^2(\sigma^6 n)^{-1}\int_{-T}^{T}|t|^5\mathrm{e}^{-\frac{3}{10}t^2}\,\mathrm{d}t\\
&\leqslant (125/54)\big[(K+2)\vee\sqrt{2\beta_3}\big]^2\beta_3^2(\sigma^6 n)^{-1}\int_0^{\infty}u^2\mathrm{e}^{-u}\,\mathrm{d}u\\
&= (125/27)\big[(K+2)\vee\sqrt{2\beta_3}\big]^2\beta_3^2(\sigma^6 n)^{-1},
\end{aligned}$$

(16) integral was evaluated

$$I_1\leqslant 5\big((K+2)\vee\sqrt{2\beta_3}\big)^2\frac{\beta_3^2}{\sigma^6 n}. \tag{18}$$

Further, since

$$\begin{aligned}
&\left|f_n(t) - \mathrm{e}^{-\frac{1}{2}t^2}\right| \\
&\leqslant \left|\exp\left\{-\frac{1}{2}t^2 - \theta(1/4)\left[(K+2)\vee\sqrt{2\beta_3}\right]\beta_3(\sigma^3\sqrt{n})^{-1}|t|^3\right\} - \exp\left\{-\frac{1}{2}t^2\right\}\right| \\
&\leqslant (1/4)\left[(K+2)\vee\sqrt{2\beta_3}\right]\beta_3(\sigma^3\sqrt{n})^{-1}|t|^3\mathrm{e}^{-\frac{3}{10}t^2},
\end{aligned}$$

so

$$\begin{aligned}
I_{2.1} &= \int_{T<|t|\leqslant T_1}\left|t^{-1}\Big(f_n(t) - \exp\left\{-\frac{1}{2}t^2\right\}\Big)\right|\mathrm{d}t \\
&\leqslant (125/36)\left[(K+2)\vee\sqrt{2\beta_3}\right]^2\beta_3^2(\sigma^6 n)^{-1}\int_0^\infty u\mathrm{e}^{-u}\,\mathrm{d}u \\
&= (125/36)\left[(K+2)\vee\sqrt{2\beta_3}\right]^2\beta_3^2(\sigma^6 n)^{-1}.
\end{aligned}$$

Moreover

$$\begin{aligned}
I_{2.2} &= \int_{T<|t|\leqslant T_1}\left|t^{-1}\gamma_3(it)^3(6\sigma^3\sqrt{n})^{-1}\mathrm{e}^{-\frac{1}{2}t^2}\right|\mathrm{d}t \\
&\leqslant (5/4)\left[(K+2)\vee\sqrt{2\beta_3}\right]^2\beta_3^2(\sigma^6 n)^{-1}\int_0^\infty u\mathrm{e}^{-u}\,\mathrm{d}u \\
&= (5/4)\left[(K+2)\vee\sqrt{2\beta_3}\right]^2(\sigma^6 n)^{-1}.
\end{aligned}$$

Thereby, inserting $I_{2.1}$, $I_{2.2}$to (17) gives

$$I_2 \leqslant (85/18)\big((K+2)\vee\sqrt{2\beta_3}\big)^2\frac{\beta_3^2}{\sigma^6 n}. \tag{19}$$

Consequently, according to (15), (18) and (19) this theorem was proofed.

Now consider lattice independent and identically distributed r.v.

THEOREM 2.3. *If independent identically distributed r.v.* $\{X_k, k = 1, 2, \ldots\}$ *are lattice, having* (1) *absolute finite moments, and the* (2) *condition is satisfied, then*

$$\sup_x\left|F_n(x) - G_n^*(x)\right| \leqslant 4\big((K+2)\vee\sqrt{2\beta_3}\big)^2\frac{(1\vee h)^4\beta_3^2}{n\sigma^6}, \tag{20}$$

here $(K+2)\vee\sqrt{2\beta_3} = \max\{K+2, \sqrt{2\beta_3}\}$.

Proof. This proof is almost similar to the 2.2 theorem's proof.

Here we have to use (9) smoothing inequality, with T and T_1, described in the 2.2 theorem. Besides

$$I_1 = \int_{|t|\leqslant T}\left|t^{-1}\big(f_n(t) - g_n^*(t)\big)\right|\mathrm{d}t \leqslant I_{1.1} + I_{1.2} + I_{1.3} \tag{21}$$

and

$$I_2 = \int_{T<|t|\leqslant T_1} \left| t^{-1}\big(f_n(t) - g_n^*(t)\big)\right| \mathrm{d}t \leqslant I_{2.1} + I_{2.2} + I_{2.3}. \tag{22}$$

Similarly to the 2.2 theorem's proof, we obtained

$$I_{1.1} + I_{1.2} \leqslant 5\big((K+2) \vee \sqrt{2\beta_3}\big)^2 \frac{\beta_3^2}{\sigma^6 n}.$$

It is easy to evaluate, that

$$I_{1.3} = I_{2.3} \leqslant h(2\pi\sigma\sqrt{n})^{-1} \sum_{v=1}^{\infty} \frac{1}{v} \exp\big\{-2\pi^2 v^2 \sigma^2 n h^{-2}\big\} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{t^2}{2}}\, \mathrm{d}t$$

$$\leqslant (\sqrt{2\pi}/8\pi^4\sqrt{\mathrm{e}})\big[(K+2) \vee \sqrt{2\beta_3}\big]^2 \frac{(1\vee h)^4 \beta_3^2}{\sigma^6 n}.$$

Since

$$I_{2.1} + I_{2.2} \leqslant (85/18)\big((K+2) \vee \sqrt{2\beta_3}\big)^2 \frac{\beta_3^2}{\sigma^6 n},$$

so inserting obtained integrals to (21), (22) and according to (9) gives (20).

Consequently, the error estimate of the normal approximation in case of the Bernoulli's distribution is

$$\sup_x \big|F_n(x) - G_n^*(x)\big| \leqslant 36 \frac{(p^2+q^2)^2}{npq}, \tag{23}$$

here r.v. achieves values 0 and 1 with positive probabilities $0 < p < 1$, $q = 1 - p$. Besides $\mathbf{E}X_1 = p$, $\mathbf{D}X_1 = pq$, $\beta_3 = pq(p^2+q^2)$.

References

1. Л. Саулис, В. Статулявичюс, *Предельные теоремы о больших уклонениях*, Мокслас (1989).
2. V.A. Statulevičius, On large devations, *Z. Wahrscheinlichkeitstheorie Verw.*, **6**, 133–144 (1966).
3. Б.В. Гнеденко, А.Н. Колмогоров, *Предельные распределения для сумм независимых случайных величин*, Наука (1949).
4. Р.Н. Бхаттачария, Р. Ранга Рао, *Аппроксимация нормальным распределением и асимптотические разложения*, Наука (1982).
5. И.А. Ибрагимов, Ю.В. Линник, *Независимые и стационарно связанные величины*, Наука (1965).

REZIUMĖ

A. Kasparavičiūtė, K. Padvelskis. Paklaidos įvertis centrinėje ribinėje teoremoje

Šiame darbe gaunamas, nepriklausomų vienodai pasiskirsčiusių atsitiktinių dydžių $\{X_k, k = 1, 2, \ldots\}$ centruotos ir normuotos sumos $S_n = \sum_{k=1}^n X_k$ skirstinio $F_n(x) = \mathbf{P}(Z_n < x)$ aproksimacijos normaliuoju dėsniu, su vienu Čebyšovo asimptotinio skleidinio nariu, tikslus įvertis.

Raktiniai žodžiai: Centrinė ribinė teorema, normalusis skirstinys, liekamojo nario įvertis, kumuliantai.