An improvement of the Berry–Esseen inequality with applications to Poisson and mixed Poisson random sums*

Victor Korolev† and Irina Shevtsova‡

15 December 2009

Abstract.

By a modification of the method that was applied in (Korolev and Shevtsova, 2009), here the inequalities

$$\rho(F_n,\Phi)\leqslant\frac{0.335789(\beta^3+0.425)}{\sqrt{n}}$$

and

$$\rho(F_n,\Phi)\leqslant\frac{0.3051(\beta^3+1)}{\sqrt{n}}$$

are proved for the uniform distance $\rho(F_n,\Phi)$ between the standard normal distribution function Φ and the distribution function F_n of the normalized sum of an arbitrary number $n\geqslant 1$ of independent identically distributed random variables with zero mean, unit variance and finite third absolute moment β^3. The first of these inequalities sharpens the best known version of the classical Berry–Esseen inequality since $0.335789(\beta^3+0.425)\leqslant 0.335789(1+0.425)\beta^3<0.4785\beta^3$ by virtue of the condition $\beta^3\geqslant 1$, and 0.4785 is the best known upper estimate of the absolute constant in the classical Berry–Esseen inequality. The second inequality is applied to lowering the upper estimate of the absolute constant in the analog of the Berry–Esseen inequality for Poisson random sums to 0.3051 which is strictly less than the least possible value of the absolute constant in the classical Berry–Esseen inequality. As a corollary, the estimates of the rate of convergence in limit theorems for compound mixed Poisson distributions are refined.

Key words: Central limit theorem, Berry–Esseen inequality, smoothing inequality, Poisson random sum, mixed Poisson distribution

*Research supported by the Russian Foundation for Basic Research, projects 08-01-00563, 08-01-00567, 08-07-00152 and 09-07-12032-ofi-m, and also by the Agency for Education of Russian Federation, state contracts P-1181 and P-958.

†Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, Moscow State University; Institute for Informatics Problems, Russian Academy of Sciences, vkorolev@comtv.ru

‡Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, Moscow State University, ishevtsova@cs.msu.su

1 Introduction

By $\mathcal{F}_3$ we will denote the set of distribution functions with zero first moment, unit second moment and finite third absolute moment β^3. Let $X_1, X_2, \ldots$ be independent random variables with common distribution function $F \in \mathcal{F}_3$ defined on a probability space $(\Omega, \mathcal{A}, \mathsf{P})$. Denote

$$F_n(x) = F^{*n}(x\sqrt{n}) = \mathsf{P}\left(\frac{X_1 + \ldots + X_n}{\sqrt{n}} < x\right),$$

$$\Phi(x) = \int_{-\infty}^{x} \phi(t)dt, \quad \phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \quad x \in \mathbb{R}.$$

The classical Berry–Esseen theorem states that there exists a finite positive absolute constant C_0 which guarantees the validity of the inequality

$$\rho(F_n, \Phi) \equiv \sup_x |F_n(x) - \Phi(x)| \leqslant C_0 \frac{\beta^3}{\sqrt{n}} \tag{1}$$

for all $n \geqslant 1$ and any $F \in \mathcal{F}_3$ (Berry, 1941), (Esseen, 1942). The problem of establishing the best value of the constant C_0 in inequality (1) is very important from the point of view of practical estimation of the accuracy of the normal approximation for the distribution functions of random variables which may be assumed to have the structure of a sum of independent random summands.

This problem has a long history and is very rich in deep and interesting results. Upper estimates for C_0 were considered in very many papers. Here we will not repeat a detailed history of the efforts to lower the upper estimates of C_0 from the original works of A. Berry (Berry, 1941) and C.-G. Esseen (Esseen, 1942) to the papers of I. S. Shiganov (Shiganov, 1982), (Shiganov, 1986) presented in (Korolev and Shevtsova, 2009). We will restrict ourselves only to an outline of the recent history of the subject.

After some lull that lasted more than twenty years, recently the interest to the problem of improving the Berry–Esseen inequality rose again and resulted in very interesting and in some sense path-clearing works. In 2006 I. G. Shevtsova improved Shiganov's upper estimate by approximately 0.06 and obtained the estimate $C_0 \leqslant 0.7056$ (Shevtsova, 2006). In 2008 she sharpened this estimate to $C_0 \leqslant 0.7005$ (Shevtsova, 2008). In 2009 the competition for improving the constant became especially keen. On 8 June, 2009 I. S. Tyurin submitted his paper (Tyurin, 2009a) to the «Theory of Probability and Its Applications». That paper, along with other results, contained the estimate $C_0 \leqslant 0.5894$. Two days later the summary of those results was submitted to «Doklady Akademii Nauk» (translated into English as «Doklady Mathematics») (Tyurin, 2009b). Independently, on 14 September, 2009 V. Yu. Korolev and I. G. Shevtsova submitted their paper (Korolev and Shevtsova, 2009) to the «Theory of Probability and Its Applications». In that paper the inequality

$$\rho(F_n, \Phi) \leqslant \frac{0.34445(\beta^3 + 0.489)}{\sqrt{n}}, \quad n \geqslant 1, \tag{2}$$

was proved which holds for any distribution $F \in \mathcal{F}_3$ yielding the estimate $C_0 \leqslant 0.5129$ by virtue of the condition $\beta^3 \geqslant 1$. Finally, on 17 November, 2009 the paper (Tyurin, 2009c) was submitted to the «Russian Mathematical Surveys» (its English version (Tyurin, 2009d) appeared on 3 December, 2009 on arXiv:0912.0726v1). In this paper the estimate $C_0 \leqslant 0.4785$ is proved. So, the best known upper estimate of the absolute constant C_0 in the classical Berry–Esseen inequality (1) is $C_0 \leqslant 0.4785$ (Tyurin, 2009c).

On the other hand, in 1956 C.-G. Esseen showed that $C_0 \geqslant C_E$ where

$$C_E = \frac{\sqrt{10}+3}{6\sqrt{2\pi}} = 0.409732...$$

(Esseen, 1956). In 1967 V. M. Zolotarev put forward the hypothesis that in (1) $C_0 = C_E$ (Zolotarev, 1967a), (Zolotarev, 1967b). However, up till now this hypothesis has been neither proved nor rejected.

To prove (2) we used an observation that from inequality (1) it obviously follows that for any $k \geqslant 0$ there exists a finite positive absolute constant C_k which guarantees the validity of the inequality

$$\rho(F_n, \Phi) \leqslant C_k \frac{\beta^3 + k}{\sqrt{n}} \tag{3}$$

for all $n \geqslant 1$ and $F \in \mathcal{F}_3$ (for example, inequality (3) trivially holds with $C_k = C_0$).

Following the lines of the reasoning we used in (Korolev and Shevtsova, 2009) to prove (2), with the only change in the way of estimation of the difference between characteristic functions in the neighborhood of zero (see lemma 2 below), in this paper we will demonstrate a special method of numerical estimation of C_k in (3). This method yields two special values of k: $k = k_0$ and $k = 1$. The first value, k_0, minimizes the upper estimate of $C_k(1+k)$ yielding the best (within the method under consideration) upper estimate of C_0 in (1) since

$$C_0 \leqslant \min_{k \geqslant 0} C_k(1+k)$$

by virue of the condition $\beta^3 \geqslant 1$. At the same time the second value, $k = 1$, minimizes C_k in (3). As we will see, $k = 1$ plays the main role in improving the absolute constant in the analog of the Berry–Esseen inequality for Poisson and mixed Poisson random sums.

Inequality (3) with $k = k_0$ and $k = 1$ is an improvement of the inequality

$$\rho(F_n, \Phi) \leqslant 0.3450 \frac{\beta^3 + 1}{\sqrt{n}}$$

we proved in (Korolev and Shevtsova, 2010a). In (Korolev and Shevtsova, 2010b) this inequality was applied to sharpening the analog of the Berry–Esseen inequality for Poisson random sums and it was for the first time demonstrated that the absolute constant in this analog can be made strictly less than that in the classical Berry–Esseen inequality.

In the papers (Shevtsova, 2010a) and (Korolev and Shevtsova, 2010a) it was shown that the constant C_k in (3) cannot be made less than the so-called lower asymptotically exact constant in the central limit theorem, that is,

$$C_k \geqslant \frac{2}{3\sqrt{2\pi}} = 0.2659...,$$

so that the gaps between the least possible value of the constant C_k and its upper estimates given in theorems 1 and 2 below are rather small and do not exceed 0.07 and

0.035, respectively, which is important from the point of view of practical applications of inequalities (6) and (7).

Our investigations were to a great extent motivated by a series of results of Håkan Prawitz and Vladimir Zolotarev outlined below.

First, since estimates of the accuracy of the normal approximation for distributions of sums of independent random variables are traditionally constructed with the use of the so-called smoothing inequalities which estimate the (uniform) distance between the pre-limit distribution function of the standardized sum of independent random variables and the limit standard normal distribution function through some integral of the (weighted) absolute value of the difference between the corresponding characteristic functions, the shape of the dependence of the final estimate on the moments of summands is fully determined by the shape of dependence of the majorant of characteristic functions on these moments. In (Prawitz, 1973) the following result was presented. Let $f(t)$ be the characteristic function corresponding to the distribution function $F \in \mathcal{F}_3$. Denote

$$\varkappa = \sup_{x>0} \frac{|\cos x - 1 + x^2/2|}{x^3} = 0.09916191...$$

and let $\theta_0 = 3.99589567...$ be the unique root of the equation

$$3(1-\cos\theta) - \theta\sin\theta - \theta^2/2 = 0,$$

lying in the interval $(\pi, 2\pi)$. Then

$$|f(t)| \leqslant \begin{cases} 1 - \dfrac{t^2}{2} + \varkappa\,(\beta_3+1)\,|t|^3, & |t| \leqslant \dfrac{\theta_0}{(\beta_3+1)}, \\ 1 - \dfrac{1-\cos\big((\beta_3+1)t\big)}{(\beta_3+1)^2}, & \theta_0 \leqslant (\beta_3+1)|t| \leqslant 2\pi, \\ 1, & |t| \geqslant \dfrac{2\pi}{(\beta_3+1)}. \end{cases}$$

As is easily seen, the majorant for $|f(t)|$ established by this inequality depends on β_3 through the function $\psi(\beta_3) = \beta_3 + 1$. This is the first hint at that the final estimate for $\rho(F_n, \Phi)$ should also depend on β_3 through the function $\psi(\beta_3) = \beta_3 + 1$.

Second, in (Prawitz, 1975b) H. Prawitz announced an inequality with unusual structure

$$\rho(F_n, \Phi) \leqslant \frac{2}{3\sqrt{2\pi}} \cdot \frac{\beta_3}{\sqrt{n-1}} + \frac{1}{2\sqrt{2\pi(n-1)}} + \frac{c_1(\beta_3)^2 + c_2\beta_3 + c_3}{n-1}, \quad n \geqslant 2,\ F \in \mathcal{F}_3, \quad (4)$$

where c_1, c_2 and c_3 are some finite positive constants. In the same paper he suggested that the coefficient

$$\frac{2}{3\sqrt{2\pi}} = 0.2659...$$

at $\beta_3/\sqrt{n-1}$ cannot be made smaller. Probably, H. Prawitz intended to publish the strict proof of (4) in the second part of his work which, unfortunately, for some reasons remained unpublished (the title of (Prawitz, 1975b) contains the Roman number I indicating the assumed continuation).

This Prawitz' inequality (4) seemed to have bepuzzled some specialists in limit theorems of probability theory. In particular, it was bypassed in the well-known books (Petrov, 1987), (Zolotarev, 1997) (in both of these books there is even no reference to any of Prawitz' works). Only in the book (Petrov, 1995) there appears a reference to the paper (Prawitz, 1975a) dealing with some estimates for characteristic functions, but the paper (Prawitz, 1975b) containing inequality (4) is again ignored. In Mathematical Reviews (Dunnage, 1977) there is only a fuzzy remark concerning «some improvements for identically distributed summands». Probably, this attitude of some specialists to inequality (4) is caused by that at first sight this inequality contradicts the Esseen's result that $C_0 \geqslant C_E$ cited above, since

$$\frac{2}{3\sqrt{2\pi}} < \frac{\sqrt{10}+3}{6\sqrt{2\pi}}.$$

However, a thorough analysis of the published part of Prawitz' work convinces that inequality (4) is valid. A strict proof of a similar inequality for not necessarily identically distributed summands with the third term being $O\big((\beta_3/\sqrt{n})^{5/3}\big)$ was given by V. Bentkus (Bentkus, 1991), (Bentkus, 1994) (for identically distributed summands, the result of Benkus is slightly worse than (4) where the third term is $O\big((\beta_3/\sqrt{n})^2\big)$).

Inequality (4) has a very interesting structure: from the main term of order $O(n^{-1/2})$ of the estimate of the accuracy of the normal approximation a summand of the form $1/\sqrt{n}$ is separated. This summand may be considerably less than the Lyapunov fraction $\beta_3/\sqrt{n}$. Moreover, in the double array scheme it may happen so that even if the Lyapunov condition $\beta_3/\sqrt{n} \to 0$ holds, the quantity $\beta_3 = \beta_3(n)$ may infinitely increase as $n \to \infty$ so that the summand of the form $n^{-1/2}$ is infinitesimal with a higher order of smallness than the Lyapunov fraction $\beta_3(n)/\sqrt{n}$. Thus, inequality (4) is the second hint at that in a reasonable estimate of $\rho(F_n, \Phi)$ depending on β_3 the term of order $O(n^{-1/2})$ should be split into two summands of the form $\beta_3/\sqrt{n}$ and $1/\sqrt{n}$ respectively.

By the way, speaking of the history of inequality (4), it has to be noted that actually it is a further development of the inequality

$$\rho(F_n, \Phi) \leqslant \frac{0.32\beta_3 + 0.25}{\sqrt{n-2}}, \quad n \geqslant 3, \tag{5}$$

which holds under the condition $\sqrt{n-1} \geqslant 3.9(\beta_3 + 1)$. The proof of (5) was given by H. Prawitz in his lecture on 16 June, 1972 at the Summer School of the Swedish Statistical Society in Löttorp (Prawitz, 1972a).

So, the final shape of inequality (3) was prompted by the works of H. Prawitz mentioned above. As this is so, the main role goes to the problem of a proper estimation of the constant C_k. To solve this problem we use a method which is a further development of the ideas of V. Zolotarev presented in (Zolotarev, 1965), (Zolotarev, 1966), (Zolotarev, 1967a) and (Zolotarev, 1967b). This method will be described in detail below.

The paper is organized as follows. In Section 2 the basic results are proved. Namely, here we prove inequality (3) with $k = k_0 = 0.425$ (theorem 1) and with $k = 1$ (theorem 2). In Section 3 theorem 2 is applied to sharpening the analog of the Berry–Esseen inequality for Poisson random sums. We show that despite a prevalent opinion that the absolute constant in this inequality should not be less than the absolute constant in the classical Berry–Esseen inequality, as a matter of fact this is not so and the constant in the Berry–Esseen inequality for Poisson random sums does not exceed 0.3051, which is, as it has been already mentioned, strictly less than the least possible value $C_E \approx 0.4097$ of the constant C_0 in (1). Finally, in Sections 4 and 5 the result of Section 3 is used for improving the estimates of the rate of convergence of compound mixed Poisson distributions with zero and non-zero means to scale and location mixtures of normal laws, respectively.

2 The basic results

2.1 Formulations and discussion

Practical calculations show that *under the algorithm we use for the estimation of* C_k (see Section 3) the resulting majorant of the constant C_k decreases as k increases from 0 to 1. At the same time for $0 \leqslant k \leqslant k_0 \approx 0.425$ the obtained estimates of $C_k(1+k)$ remain constant, and for $k > k_0$ they begin to increase although in the interval $k_0 < k < 1$ the obtained estimate of C_k decreases. Thus, we can present two computationally optimal values of k in (3): $k_0 = 0.425$ and $k_1 = 1$. The first of them delivers the minimum value to the upper estimate of $C_k(1+k)$, thus solving the problem of estimation of C_0 in (1), whereas the second, maximin, minimizes the estimate of C_k in (3).

The use of $k = k_0$ in (3) gives the following result.

THEOREM 1. *For all $n \geqslant 1$ and all distributions with zero mean, unit variance and finite third absolute moment β^3 we have the inequality*

$$\rho(F_n, \Phi) \leqslant \frac{0.335789(\beta^3 + 0.425)}{\sqrt{n}}. \tag{6}$$

REMARK 1. Under the conditions imposed on the moments of the random variable X_1 we always have $\beta^3 \geqslant 1$. Therefore,

$$0.335789(\beta^3 + 0.425) \leqslant 0.335789(1 + 0.425)\beta^3 < 0.4785\beta^3.$$

Hence, inequality (6) is always sharper than the classical Berry–Esseen inequality (1) with the best known constant $C_0 = 0.4785$ for all possible values of β^3, although the same prior information concerning the distribution F is required for its validity (namely, only the value of the third absolute moment β^3).

REMARK 2. Inequality (6) is an «unconditional» variant of the «conditional» Prawitz inequality (5) and is a practically computable analog of inequality (4) with a slightly (approximately by 0.07) worse first coefficient and a slightly better (approximately by 0.05) second coefficient, but without the third summand that contains unknown constants.

REMARK 3. Even if the hypothesis of V. M. Zolotarev that $C_0 = C_E = 0.4097...$ in (1) (see (Zolotarev, 1967a), (Zolotarev, 1967b)) turns out to be true, then, due to that $\beta^3 \geqslant 1$, inequality (6) will be sharper than the classical Berry–Esseen inequality (1) for $\beta^3 \geqslant 1.93$.

The use of $k = 1$ in (3) yields the following result.

THEOREM 2. *For all $n \geqslant 1$ and all distributions with zero mean, unit variance and finite third absolute moment β^3 we have the inequality*

$$\rho(F_n, \Phi) \leqslant \frac{0.3051(\beta^3 + 1)}{\sqrt{n}}. \tag{7}$$

REMARK 4. Inequality (7) is another «unconditional» variant of the «conditional» Prawitz inequality (5). Moreover, the first coefficient in (7) is less than that in (5) by approximately 0.02 whereas the second coefficient in (7) is greater than that in (5) by approximately 0.05.

2.2 Proofs of basic results

2.2.1 Auxiliary statements

As we have already mentioned above, to prove theorem 1 we will follow the lines of the approach proposed and developed by V. M. Zolotarev in his works (Zolotarev, 1965), (Zolotarev, 1966) and (Zolotarev, 1967). This approach is based on the application of smoothing inequalities which make it possible to estimate the distance between distribution functions via the distances between the corresponding characteristic functions. Within this approach the key points are: (i) the choice of a proper smoothing inequality; (ii) the choice of a proper smoothing kernel in a smoothing inequality; (iii) the choice of proper estimates for the distance between characteristic functions; (iv) the choice of a proper computational optimization procedure.

We will describe these points one after another as they are used in the proof of theorems 1 and 2. The corresponding statements will have the form of lemmas.

We begin with the smoothing inequality. In most papers dealing with the estimation of the constant in the Berry–Esseen inequality (1) smoothing inequalities of the same type were used. This type of smoothing inequalities was introduced by V. M. Zolotarev. In the original paper (Zolotarev, 1965), just as in similar inequalities in the earlier papers of Berry (Berry, 1941) and Esseen (Esseen, 1942), the kernel was used which had a probabilistic sense, that is, which was the probability density of some symmetric probability distribution. In the paper of Van Beek (Van Beek, 1972) it was noticed that this condition is not crucial. Van Beek proposed to use symmetric kernels with alternating signs. Concurrently with (Van Beek, 1972), the paper of V. Paulauskas (Paulauskas, 1971) was published in which the original smoothing inequality of Zolotarev was generalized (and hence, sharpened) to the case of positive non-symmetric kernels. It is interesting to notice that although in the final part of the paper of Paulauskas it was noted that the smoothing inequality proved in that paper was destined, in the first place, for improving the constant in the Berry–Esseen inequality, as far as we know, unfortunately no one ever used the Paulauskas inequality for this purpose. In (Shevtsova, 2009b) a new smoothing inequality was proved which generalizes (and hence, sharpens) both Paulauskas' and Van Beek's inequalities to the case of non-symmetric kernels with alternating signs. However, all these inequalities yield worse estimates than the Prawitz smoothing inequality proved in (Prawitz, 1972b).

The characteristic function of the standardized sum $(X_1+\ldots+X_n)/\sqrt{n}$ will be denoted $f_n(t)$,

$$f_n(t) = \int\limits_{-\infty}^{\infty} e^{itx} dF_n(x), \quad t \in \mathbb{R}.$$

Also denote

$$r_n(t) = |f_n(t) - e^{-t^2/2}|.$$

LEMMA 1 (Prawitz, 1972b). *For an arbitrary distribution function F and $n \geqslant 1$ for any $0 < t_0 \leqslant 1$ and $T > 0$ we have the inequality*

$$\rho(F_n, \Phi) \leqslant 2\int_0^{t_0} |K(t)| r_n(Tt)\, dt + 2\int_{t_0}^1 |K(t)| \cdot |f_n(Tt)| dt +$$

$$+2\int_0^{t_0} \left| K(t) - \frac{i}{2\pi t} \right| e^{-T^2t^2/2} dt + \frac{1}{\pi}\int_{t_0}^{\infty} e^{-T^2t^2/2} \frac{dt}{t},$$

where

$$K(t) = \frac{1}{2}(1-|t|) + \frac{i}{2}\left[(1-|t|)\cot \pi t + \frac{\mathrm{sign} t}{\pi}\right], \quad -1 \leqslant t \leqslant 1. \tag{8}$$

REMARK 5. In (Vaaler, 1985) a proof of a result similar to the Prawitz inequality stated by lemma 1 was given by a techniques different from that used in (Prawitz, 1972b) and it was also proved that the kernel $K(t)$ defined by (8) is in some sense optimal.

Now consider the estimates of the characteristic functions appearing in lemma 1. For $\varepsilon > 0$ set

$$\chi(t, \varepsilon) = \begin{cases} t^2/2 - \varkappa\varepsilon|t|^3, & |t| \leqslant \theta_0/\varepsilon, \\ \dfrac{1 - \cos \varepsilon t}{\varepsilon^2}, & \theta_0 < \varepsilon|t| \leqslant 2\pi, \\ 0, & |t| > 2\pi/\varepsilon, \end{cases} \tag{9}$$

where $\theta_0 = 3.99589567...$ is the unique root of the equation

$$\theta^2 + 2\theta \sin \theta + 6(\cos \theta - 1) = 0, \quad \pi \leqslant \theta \leqslant 2\pi, \tag{10}$$

$$\varkappa \equiv \sup_{x>0} \frac{|\cos x - 1 + x^2/2|}{x^3} = \frac{\cos x - 1 + x^2/2}{x^3}\bigg|_{x=\theta_0} = 0.09916191... \tag{11}$$

It can easily be made sure that the function $\chi(t, \varepsilon)$ monotonically decreases in $\varepsilon > 0$ for any fixed $t \in \mathbb{R}$.

The Lyapunov fraction will be denoted $\ell = \beta^3/\sqrt{n}$. In addition, denote

$$\ell_n = \ell + 1/\sqrt{n}.$$

LEMMA 2. *For any $F \in \mathcal{F}_3$, $n \geqslant 1$ and $t \in \mathbb{R}$ the following estimates take place:*

$$|f_n(t)| \leqslant \left[1 - \frac{2}{n}\chi(t, \ell_n)\right]^{n/2} \equiv f_1(t, \ell_n, n),$$

$$|f_n(t)| \leqslant \exp\{-\chi(t,\ell_n)\} \equiv f_2(t,\ell_n),$$

$$|f_n(t)| \leqslant \exp\Big\{-\frac{t^2}{2}+\varkappa\ell_n|t|^3\Big\} \equiv f_3(t,\ell_n).$$

REMARK 6. Apparently, the function $f_1(t,\ell_n,n)$ was used in the problem of numerical evaluation of the absolute constants in the estimates of the accuracy of the normal approximation for the first time in (Korolev and Shevtsova, 2009). The second and the third estimates presented in lemma 2 are due to H. Prawitz (Prawitz, 1973), (Prawitz, 1975b).

REMARK 7. Evidently, $f_1(t,\varepsilon,n) \leqslant f_2(t,\varepsilon)$ for all $n \geqslant 1$, $\varepsilon > 0$ and $t \in \mathbb{R}$. Moreover, from the result of Prawitz (Prawitz, 1973) it follows that $f_2(t,\varepsilon) \leqslant f_3(t,\varepsilon)$ for all $\varepsilon > 0$ and $t \in \mathbb{R}$, thus the sharpest estimate for $|f_n(t)|$ is given by $f_1(t,\ell_n,n)$, while the estimates $f_j(t,\ell_n)$, $j=2,3$, possess a useful property of monotonicity in ℓ_n which is very important for the computational procedure.

LEMMA 3 (Tyurin, 2009a), (Tyurin, 2009c), (Tyurin, 2009d). *For any $F \in \mathcal{F}_3$, $n \geqslant 1$ and $t \in \mathbb{R}$ we have*

$$r_n(t) \leqslant \ell e^{-t^2/2}\int_0^{|t|}\frac{u^2}{2}\,e^{u^2/2}\Big|f\Big(\frac{u}{\sqrt{n}}\Big)\Big|^{n-1}du.$$

The combination of lemmas 2 and 3 allows to obtain an estimate for the difference of the characteristic functions in the neighborhood of zero, which is sharper than all the analogous estimates used in the preceding works:

$$r_n(t) \leqslant \ell e^{-t^2/2}\int_0^{|t|}\frac{u^2}{2}\,e^{u^2/2}\Big[1-\frac{2}{n}\chi\Big(u,\ell+\frac{1}{\sqrt{n}}\Big)\Big]^{(n-1)/2}du \equiv r_1(t,\ell,n), \quad t\in\mathbb{R}.$$

From what was said above it follows that the substitution of the functions $f_j(t,\ell_n)$, $j = 2,3$, instead of $f_1(t,\ell_n,n)$ into the right-hand side of the last inequality does not make the resulting estimate less, thus, we obtain two more estimates for $r_n(t)$ which monotonically increase in ℓ:

$$r_n(t) \leqslant \ell e^{-t^2/2}\int_0^{|t|}\frac{u^2}{2}\,e^{u^2/2}\exp\Big\{-\frac{n-1}{n}\cdot\chi\Big(u,\ell+\frac{1}{\sqrt{n}}\Big)\Big\}du \equiv r_2(t,\ell,n),$$

$$r_n(t) \leqslant \ell e^{-t^2/2}\int_0^{|t|}\frac{u^2}{2}\exp\Big\{\varkappa\ell_n u^3+\frac{u^2}{2n}\Big(1-2\varkappa\ell_n u\Big)\Big\}du \equiv r_3(t,\ell,n), \quad t\in\mathbb{R},$$

(recall that $\ell_n = \ell + 1/\sqrt{n}$).

Noticing that

$$\Big|K(t)-\frac{i}{2\pi t}\Big| = \frac{1}{2}(1-t)\sqrt{1+\Big(\cot\pi t-\frac{1}{\pi t}\Big)^2}, \quad 0\leqslant t\leqslant 1,$$

we can estimate $\rho(F_n,\Phi)$ for any $n \geqslant 2$ and F with a fixed Lyapunov fraction ℓ as

$$\rho(F_n,\Phi) \leqslant 2\int_0^{t_0}|K(t)|\cdot r_1(Tt,\ell,n)\,dt + 2\int_{t_0}^{1}|K(t)|\cdot f_1(Tt,\ell+1/\sqrt{n},n)\,dt +$$

$$+\frac{1}{\pi}\int_{t_0}^{\infty} e^{-T^2t^2/2}\frac{dt}{t} + \int_0^{t_0}(1-t)\sqrt{1+\Big(\cot \pi t - \frac{1}{\pi t}\Big)^2}e^{-T^2t^2/2}dt \equiv D(\ell, n, t_0, T)$$

with arbitrary positive T and $t_0 \leqslant 1$.

The following lemma makes it possible to bound above the set of the values of n under consideration when estimating the constant C_k in inequality (3) with $0 < k \leqslant 1$.

LEMMA 4. *For any positive N, $k \leqslant 1$ and $\varepsilon > k/\sqrt{N}$ for all $t \in \mathbb{R}$ the following estimates hold:*

$$\sup_{n\geqslant N} f_j\Big(t, \varepsilon + \frac{1-k}{\sqrt{n}}\Big) = f_j\Big(t, \varepsilon + \frac{1-k}{\sqrt{N}}\Big) \equiv \widetilde{f}_{j,N}(t,\varepsilon), \quad j = 1, 2,$$

$$\sup_{n\geqslant N} r_2\Big(t, \varepsilon - \frac{k}{\sqrt{n}}, n\Big) \leqslant \varepsilon e^{-t^2/2}\int_0^{|t|}\frac{u^2}{2}\exp\Big\{\frac{u^2}{2} - \frac{N-1}{N}\chi\Big(u, \varepsilon + \frac{1-k}{\sqrt{N}}\Big)\Big\}du \equiv \widetilde{r}_{2,N}(t,\varepsilon),$$

For

$$|t| \leqslant T(N,\varepsilon) \equiv \min\Big\{N^{1/4}\varepsilon^{-1/2}, (2\varkappa\varepsilon)^{-1}\Big\}$$

we also have the estimate

$$\sup_{n\geqslant N} r_3\Big(t, \varepsilon - \frac{1}{\sqrt{n}}, n\Big) \leqslant \frac{1}{6\varkappa}\Big(e^{\varkappa\varepsilon|t|^3} - 1\Big)e^{-t^2/2} \equiv \widetilde{r}_3(t,\varepsilon).$$

P r o o f. The first two statements are trivial consequences of the monotonicity of the functions $\chi(t, \varepsilon + (1-k)/\sqrt{n})$ and $f_j(t, \varepsilon + (1-k)/\sqrt{n})$, $j = 1, 2$, with respect to $n \geqslant 1$.

To prove the third statement note that the function r_3 can be written in the form

$$r_3\Big(t, \varepsilon - \frac{1}{\sqrt{n}}, n\Big) = e^{-t^2/2}\int_0^{|t|}\frac{u^2}{2}\exp\big\{\varkappa\varepsilon u^3 + g(n,u)\big\}du,$$

where

$$g(x,u) = \ln\Big(\varepsilon - \frac{1}{\sqrt{x}}\Big) + \frac{a(u)}{x}, \ x > 0, \quad a(u) = \frac{u^2}{2}\Big(1 - 2\varkappa\varepsilon u\Big), \ u > 0.$$

Since $|t| \leqslant (2\varkappa\varepsilon)^{-1}$ under the conditions of the lemma, we have $a(u) \geqslant 0$ for all $u \leqslant |t|$. Let us establish that $g(x,u)$ monotonically increases in $x \geqslant N$ and $u \leqslant T(N,\varepsilon)$. Indeed, the derivative

$$\frac{\partial g(x,u)}{\partial x} = \frac{1}{2x(\varepsilon\sqrt{x} - 1)} - \frac{a(u)}{x^2}$$

is non-negative if and only if $x - 2a(u)\varepsilon\sqrt{x} + 2a(u) \geqslant 0$. Since $a(u) \geqslant 0$, the last condition is satisfied, if $\sqrt{x} \geqslant 2a(u)\varepsilon \equiv \varepsilon u^2(1 - 2\varkappa\varepsilon u)$, or, particulary, if $\sqrt{x} \geqslant \varepsilon u^2$. So, for all $x \geqslant N$ and $u \leqslant T(N,\varepsilon)$ with $T(N,\varepsilon)$ defined in the formulation of the lemma the function $g(x,u)$ monotonically increases in $x \geqslant N$, whence it follows that

$$\sup_{n\geqslant N} g(n,u) = \lim_{n\to\infty} g(n,u) = \ln\varepsilon, \quad 0 \leqslant u \leqslant T(N,\varepsilon),$$

and

$$\sup_{n\geqslant N} r_3\Big(t, \varepsilon - \frac{1}{\sqrt{n}}, n\Big) \leqslant \frac{1}{2}\varepsilon e^{-t^2/2}\int_0^{|t|} u^2 e^{\varkappa\varepsilon u^3}du = \frac{1}{6\varkappa}\Big(e^{\varkappa\varepsilon|t|^3} - 1\Big)e^{-t^2/2} \equiv \widetilde{r}_3(t,\varepsilon),$$

Q. E. D.

Finally, the process of computational optimization can be properly organized with the help of the following statements.

LEMMA 5 (Bhattacharya and Ranga Rao, 1976). *For any distribution F with zero mean and unit variance we have*

$$\rho(F, \Phi) \leqslant \sup_{x>0} \left(\Phi(x) - \frac{x^2}{1+x^2} \right) = 0.54093654\ldots$$

LEMMA 6. *For any $F \in \mathcal{F}_3$ and $n \geqslant 400$ such that $\beta^3 + 1 \leqslant 0.1\sqrt{n}$ the following estimate takes place*:

$$\rho(F_n, \Phi) \leqslant 0.2727 \cdot \frac{\beta^3}{\sqrt{n}} + \frac{0.2041}{\sqrt{n}}.$$

The statement of lemma 6 is a result of the algorithm described in (Prawitz, 1975b) or (Gaponova and Shevtsova, 2009).

Since the function

$$g(b) = \frac{0.2727b + 0.2041}{b+k}, \quad b \geqslant 1,$$

monotonically increases for $k > 0.2041/0.2727 = 0.74\ldots$ and monotonically decreases for $0 \leqslant k \leqslant 0.74$, we have

$$\sup_{b\geqslant 1} g(b) = \begin{cases} 0.2727, & k \geqslant 0.75, \\ 0.4768/(1+k), & k \leqslant 0.74. \end{cases}$$

Thus, from lemma 6 it follows that for all n and β^3 such that $(\beta^3 + k)/\sqrt{n} < 0.05(1+k)$ inequality (3) holds with $C_k = 0.2727$ for $k \geqslant 0.75$ and with $C_k = 0.4768/(1+k)$ for $k \leqslant 0.74$. In particular, for $k = 0.425$ we have

$$\rho(F_n, \Phi) \leqslant 0.3346 \cdot \frac{\beta^3 + 0.425}{\sqrt{n}}, \quad \text{if} \quad \frac{\beta^3 + 0.425}{\sqrt{n}} \leqslant 0.07125.$$

The lemmas presented above give the grounds for restricting the domain of the values of $\varepsilon = (\beta^3 + k)/\sqrt{n}$ by a bounded interval separated from zero (more details will be given below) and for looking for the constant C_k in the form

$$C_k = \max_{\varepsilon} C(\varepsilon), \quad C(\varepsilon) = D(\varepsilon)/\varepsilon, \quad D(\varepsilon) = \sup\left\{D(\varepsilon, n) \colon n \geqslant n_*\right\}, \tag{11}$$

where

$$D(\varepsilon, n) = \inf_{0<t_0\leqslant 1,\, T>0} D(\varepsilon - k/\sqrt{n}, n, t_0, T), \tag{12}$$

$$n_* = \max\{1, \lceil (1+k)^2/\varepsilon^2 \rceil\},$$

here $\lceil x \rceil$ is the least integer no less than x. The condition $n \geqslant n_*$ is a consequence of the inequality $\beta^3 \geqslant 1$. For the estimation of the supremum in n in the definition of $D(\varepsilon)$, lemma 4 is used for N large enough. The computation of the maximum in ε is essentially based on the property of monotonicity in ε of all the functions used for the estimation of $|f_n(t)|$ and $r_n(t)$, and hence, on the monotonicity of the function $D(\varepsilon) = \varepsilon C(\varepsilon)$. This property makes it possible to estimate $\max_\varepsilon C(\varepsilon)$ using the values of $C(\varepsilon)$ only in a finite number of points. In particular, the following statement holds.

LEMMA 7. *For all $\varepsilon_2 > \varepsilon_1 > 0$ the following inequality is true*:

$$\max_{\varepsilon_1 \leqslant \varepsilon \leqslant \varepsilon_2} C(\varepsilon) \leqslant C(\varepsilon_2) \cdot \frac{\varepsilon_2}{\varepsilon_1}.$$

2.2.2 Proof of theorem 1

Denote

$$\varepsilon = \ell + \frac{0.425}{\sqrt{n}} = \frac{\beta^3 + 0.425}{\sqrt{n}}.$$

Then for $\varepsilon \leqslant 0.07$ inequality (6) is a consequence of lemma 6, and for $\varepsilon \geqslant 1.62 \geqslant 0.541/0.335789$ it follows from lemma 5. Thus, to compute C_k the maximization with respect to ε in (11) is conducted on the interval $0.07 \leqslant \varepsilon \leqslant 1.62$. To compute the supremum with respect to $n \geqslant n_* = \lceil (1.425/\varepsilon)^2 \rceil$ we use lemma 4 with $N = 600$ for $\varepsilon \leqslant 0.1$, $N = 300$ for $0.1 < \varepsilon \leqslant 0.2$ and $N = 100$ for $\varepsilon > 0.2$. For the mentioned values of ε we have $n_*(0.07) = 415$, $n_*(0.1) = 204$, $n_*(0.2) = 51$. The maximum with respect to ε is estimated by lemma 7 and is attained in the two points: $n = 5$, $\varepsilon \approx 0.822$ ($\beta^3 \approx 1.413$, $t_0 \approx 0.385$, $T = 5.755$) and $n = 8$, $\varepsilon \approx 0.504$ ($\beta^3 = 1$, $t_0 \approx 0.293$, $T = 8.911$). Both extremal values do not exceed 0.335789, whence, theorem 1 is proved.

2.2.3 Proof of theorem 2

Denote

$$\varepsilon = \ell + \frac{1}{\sqrt{n}} = \frac{\beta^3 + 1}{\sqrt{n}}.$$

Then for $\varepsilon \leqslant 0.1$ inequality (7) is a consequence of lemma 6, and for $\varepsilon \geqslant 1.78 \geqslant 0.541/0.3051$ it follows from lemma 5. Thus, to compute C_k the maximization with respect to ε in (11) is conducted on the interval $0.1 \leqslant \varepsilon \leqslant 1.78$. To compute the supremum with respect to $n \geqslant n_* = \lceil 4/\varepsilon^2 \rceil$ we use the last statement of lemma 4 with $N = 200$ and $T(200, \varepsilon) = \min\{5.04/\varepsilon, 3.76/\sqrt{\varepsilon}\}$. It turned out, that the extremal value is attained at $n \to \infty$ and $\varepsilon \approx 0.985$ ($t_0 = 0.356$, $T = 6.147$) and it does not exceed 0.3051, Q. E. D.

3 An improvement of the analog of the Berry–Esseen inequality for Poisson random sums

3.1 The history of the problem

In this section we will use theorem 1 to improve the analog of the Berry–Esseen inequality for Poisson random sums. Let $X_1, X_2, \ldots$ be independent identically distributed random variables with

$$\mathsf{E}X_1 \equiv \mu, \quad \mathsf{D}X_1 \equiv \sigma^2 > 0 \quad \text{and} \quad \mathsf{E}|X_1|^3 \equiv \beta^3 < \infty. \tag{13}$$

Let N_λ be a random variable with the Poisson distribution with parameter $\lambda > 0$. Assume that for any $\lambda > 0$ the random variables N_λ and $X_1, X_2, \ldots$ are independent. Set

$$S_\lambda = X_1 + \ldots + X_{N_\lambda}$$

(for definiteness we assume that $S_\lambda = 0$ if $N_\lambda = 0$). Poisson random sums S_λ are very popular mathematical models of many real objects. In particular, in insurance mathematics S_λ describes the total claim size under the classical risk process in the «dynamical» case. Many examples of applied problems from various fields where Poisson random sums are encountered can be found in, say, (Gnedenko and Korolev, 1996) or (Bening and Korolev, 2002).

It is easy to see that

$$\mathsf{E}S_\lambda = \lambda\mu, \qquad \mathsf{D}S_\lambda = \lambda(\mu^2 + \sigma^2).$$

The distribution function of the standardized Poisson random sum

$$\widetilde{S}_\lambda \equiv \frac{S_\lambda - \lambda\mu}{\sqrt{\lambda(\mu^2 + \sigma^2)}}$$

will be denoted $F_\lambda(x)$.

It is well known that under the conditions on the moments of the random variable X_1 given above, the so-called Berry–Esseen inequality for Poisson random sums holds: there exists an absolute positive constant $C < \infty$ such that

$$\rho(F_\lambda, \Phi) \equiv \sup_x |F_\lambda(x) - \Phi(x)| \leqslant C \frac{\beta^3}{(\mu^2 + \sigma^2)^{3/2}\sqrt{\lambda}}. \tag{14}$$

Inequality (14) has rather an interesting history. Apparently, it was first proved in (Rotar, 1972a) and was published in (Rotar, 1972b) with $C = 2.23$ (the dissertation (Rotar, 1972a) was not published whereas the paper (Rotar, 1972b) does not contain a proof of this result). Later, with the use of a traditional technique based on the Esseen smoothing inequality this estimate was proved in (von Chossy, Rappl, 1983) with $C = 2.21$ (the authors of this paper declared that $C = 3$ in the formulation of the corresponding theorem, which is, of course, true, but actually in the proof of this theorem they obtained the value $C = 2.21$).

In the paper (Michel, 1993) the property of infinite divisibility of compound Poisson distributions was used to prove that the constant in (14) is the same as that in the classical Berry–Esseen inequality. Although Shiganov's estimate $C_0 \leqslant 0.7655$ (Shiganov, 1986), had been known by that time (the original paper by Shiganov had been published in Russian even earlier, in 1982), Michel used the previous record value due to Van Beek (Van Beek, 1972) and announced in (Michel, 1993) that $C \leqslant 0.8$ in (14). Being not aware of this paper of Michel, the authors of the paper (Bening, Korolev and Shorgin, 1997) used an improved version of the Esseen smoothing inequality and obtained the estimate $C \leqslant 1.99$. As it has been already noted, the method of the proof used in (Michel, 1993) is based on the fact that if for the absolute constant C_0 in the classical Berry–Esseen inequality (1) an estimate $C_0 \leqslant M$ is known, then inequality (14) holds with $C = M$.

This circumstance was also noted by the authors of the paper (Korolev and Shorgin, 1997) in which independently of the paper (Michel, 1993) the same result was obtained, but with another currently best estimate $M = 0.7655$. As we noted in Section 1, the best known estimate of the absolute constant in the classical Berry–Esseen inequality was obtained in (Tyurin, 2009c), (Tyurin, 2009d): $C_0 \leqslant 0.4785$. Therefore, following the logics of the reasoning used in (Michel, 1993) and (Korolev and Shorgin, 1997) we can conclude that inequality (14) holds at least with $C = 0.4785$.

In this section we show that actually binding the estimate of the constant C in (14) to the estimate of the absolute constant C_0 in the classical Berry–Esseen inequality is more loose. Namely, although the best known upper estimate of C_0 is $M = 0.4785$ and moreover, although the unimprovable lower estimate of C_0 is $\approx 0.4097...$, inequality (14) actually holds with $C = 0.3051$. Thus, here we improve the result of (Korolev and Shevtsova, 2010b) where we proved inequality (14) with $C = 0.3450$.

3.2 Auxiliary results

The following lemma determines the relation between the distributions and moments of Poisson random sums and the distributions and moments of sums of a non-random number of independent summands. This lemma will be the main tool which we will use to apply the results known for the classical case, to Poisson random sums.

Here and in what follows the symbol $\stackrel{d}{=}$ will stand for the coincidence of distributions. Also denote $\nu = \lambda/n$.

LEMMA 7. *The distribution of the Poisson random sum S_λ coincides with the distribution of the sum of a non-random number n of independent identically distributed random variables whatever integer $n \geqslant 1$ is:*

$$X_1 + \ldots + X_{N_\lambda} \stackrel{d}{=} Y_{\nu,1} + \ldots + Y_{\nu,n}$$

where for each n the random variables $Y_{\nu,1}, \ldots, Y_{\nu,n}$ are independent and identically distributed. Moreover, if the random variable X_1 satisfies conditions (13), *then for the moments of the random variable $Y_{\nu,1}$ the following relations hold*:

$$\mathsf{E}\, Y_{\nu,1} = \mu\nu, \quad \mathsf{D}\, Y_{\nu,1} = (\mu^2 + \sigma^2)\nu,$$

$$\mathsf{E}\, |Y_{\nu,1} - \mu\nu|^3 \leqslant \nu\beta^3(1 + 40\nu) \quad \text{for} \quad n \geqslant \lambda.$$

P r o o f. The proof is based on the property of infinite divisibility of a compound Poisson distribution which implies that for any integer $n \geqslant 1$ the characteristic function $f_{S_\lambda}(t)$ of the Poisson random sum S_λ can be represented as

$$f_{S_\lambda}(t) = \exp\big\{\lambda(f(t) - 1)\big\} = \big[\exp\big\{\nu(f(t) - 1)\big\}\big]^n \equiv \big[f_{Y_{\nu,1}}(t)\big]^n,$$

where $f_{Y_{\nu,1}}$ is the characteristic function of the random variable $Y_{\nu,1}$. Hence, the distribution of each of the summands $Y_{\nu,1}, \ldots, Y_{\nu,n}$ coincides with the distribution of the Poisson random sum of the original random variables:

$$Y_{\nu,k} \stackrel{d}{=} X_1 + \ldots + X_{N_\nu}, \quad k = 1, \ldots, n,$$

where N_ν is the Poisson-distributed random variable with parameter ν independent of the sequence $X_1, X_2, \ldots$ Hence we directly obtain the relations for the first and the second moments of the random variables $Y_{\nu,1}$ and X_1. Let us prove the relation for the third absolute moments. By the formula of total probability we have

$$\mathsf{E}\,|Y_{\nu,1} - \mu\nu|^3 \leqslant e^{-\nu}\Big(\nu^3|\mu|^3 + \nu\mathsf{E}\,|X_1 - \mu\nu|^3 + \sum_{k=2}^{\infty}\frac{\nu^k}{k!}\mathsf{E}\,|X_1 + \ldots + X_k - \mu\nu|^3\Big).$$

Consider the second and the third summands on the right-hand side separately. For this purpose without loss of generality we will assume that $n \geqslant \lambda$, that is, $\nu \leqslant 1$. By virtue of the Minkowski inequality we have

$$\big(\mathsf{E}\,|X_1 - \mu\nu|^3\big)^{1/3} \leqslant (\beta^3)^{1/3} + |\mu|\nu = (\beta^3)^{1/3}\left(1 + \frac{|\mu|\nu}{(\beta^3)^{1/3}}\right).$$

Since $\nu \leqslant 1$ and the ratio $|\mu|/(\beta^3)^{1/3}$ does not exceed 1 by virtue of the Lyapunov inequality, we obtain

$$\mathsf{E}\,|X_1 - \mu\nu|^3 \leqslant \beta^3(1+\nu)^3 \leqslant \beta^3(1 + 7\nu).$$

To estimate the third summand notice that the Lyapunov inequality yields

$$\Big|\sum_{i=1}^{k} x_i\Big|^r \leqslant k^{r-1}\sum_{i=1}^{k}|x_i|^r, \quad x_i \in \mathbb{R},\ i = 1, \ldots, k,\ r \geqslant 1,$$

(see, e. g., (Bhattacharya and Ranga Rao, 1976)). With $r = 3$, this inequality implies

$$\mathsf{E}\,|X_1 + \ldots + X_k - \mu\nu|^3 \leqslant \mathsf{E}\,(|X_1| + \ldots + |X_k| + |\mu|\nu)^3 \leqslant$$

$$\leqslant (k+1)^2(k\beta^3 + (|\mu|\nu)^3) \leqslant \beta^3(k+1)^3$$

(here we took into account that $|\mu|^3 \leqslant \beta^3$ and $\nu \leqslant 1$). Thus,

$$\mathsf{E}\,|Y_{\nu,1} - \mu\nu|^3 \leqslant \nu^3|\mu|^3 + \nu\mathsf{E}\,|X_1 - \mu\nu|^3 + \sum_{k=2}^{\infty}\frac{\nu^k}{k!}\mathsf{E}\,|X_1 + \ldots + X_k - \mu\nu|^3 \leqslant$$

$$\leqslant \nu\beta^3\big[1 + (8 + K)\nu\big]$$

where

$$K = \sum_{k=2}^{\infty}\frac{(k+1)^3}{k!} = 15e - 9 < 32.$$

The lemma is proved.

Corollary 1. *Under conditions* (13) *the distribution of the standardized Poisson random sum $\widetilde{S}_\lambda$ coincides with the distribution of the normalized non-random sum of n random variables whatever integer $n \geqslant 1$ is*:

$$\widetilde{S}_\lambda \stackrel{d}{=} \frac{1}{\sqrt{n}}\sum_{k=1}^{n} Z_{\nu,k}$$

where for each n the random variables $Z_{\nu,1},\ldots,Z_{\nu,n}$ are independent and identically distributed. Moreover, these random variables have zero expectation, unit variance and for all $n\geqslant\lambda$ their third absolute moment satisfies the inequality

$$\mathsf{E}\,|Z_{\nu,1}|^3\leqslant\frac{\beta^3(1+40\nu)\sqrt{n}}{(\mu^2+\sigma^2)^{3/2}\sqrt{\lambda}}. \tag{15}$$

P r o o f. According to lemma 7 for any n we have the representation

$$\widetilde{S}_\lambda=\frac{S_\lambda-\lambda\mu}{\sqrt{\lambda(\mu^2+\sigma^2)}}\ \stackrel{d}{=}\ \frac{Y_{\nu,1}+\ldots+Y_{\nu,n}-n\mu\nu}{\sqrt{(\mu^2+\sigma^2)n\nu}}\ \equiv\ \frac{1}{\sqrt{n}}\sum_{k=1}^{n}Z_{\nu,k},$$

in which the random variables

$$Z_{\nu,k}\equiv\frac{Y_{\nu,k}-\mu\nu}{\sqrt{\nu}}=\frac{Y_{\nu,k}-\mathsf{E}\,Y_{\nu,k}}{\sqrt{\mathsf{D}\,Y_{\nu,k}}}$$

are independent, identically distributed, have zero expectation and, unit variance. Moreover, by virtue of the same lemma for all $n\geqslant\lambda$ we have the relation

$$\mathsf{E}\,|Z_{\nu,1}|^3=\frac{\mathsf{E}\,|Y_{\nu,1}-\mathsf{E}\,Y_{\nu,1}|^3}{(\mathsf{D}\,Y_{\nu,1})^{3/2}}\leqslant\frac{\beta^3(1+40\nu)}{(\mu^2+\sigma^2)^{3/2}\nu^{1/2}}=\frac{\beta^3(1+40\nu)\sqrt{n}}{(\mu^2+\sigma^2)^{3/2}\sqrt{\lambda}}.$$

The corollary is proved.

3.3 Main result

THEOREM 2. *Under conditions* (13) *for any $\lambda>0$ we have the inequality*

$$\rho(F_\lambda,\Phi)\leqslant\frac{0.3051\beta^3}{(\mu^2+\sigma^2)^{3/2}\sqrt{\lambda}}.$$

P r o o f. From lemma 7 and corollary 1 it follows that for any integer $n\geqslant1$

$$\rho(F_\lambda,\Phi)=\sup_x\left|\mathsf{P}\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}Z_{\nu,k}<x\right)-\Phi(x)\right|.$$

Hence, by theorem 1 for an arbitrary integer $n\geqslant1$ we have

$$\rho(F_\lambda,\Phi)\leqslant0.3051\frac{\mathsf{E}|Z_{\nu,1}|^3}{\sqrt{n}}+\frac{0.3051}{\sqrt{n}}. \tag{16}$$

Since $n\geqslant1$ is arbitrary, we can assume that $n\geqslant\lambda$, making it possible to use estimate (15) for the specified n and, in the continuation of (16), to obtain the inequality

$$\rho(F_\lambda,\Phi)\leqslant0.3051\frac{\beta^3(1+40\lambda/n)}{(\mu^2+\sigma^2)^{3/2}\sqrt{\lambda}}+\frac{0.3051}{\sqrt{n}}.$$

Since here $n\geqslant\lambda$ is arbitrary, letting $n\to\infty$ we finally obtain

$$\rho(F_\lambda,\Phi)\leqslant\lim_{n\to\infty}\left[0.3051\frac{\beta^3(1+40\lambda/n)}{(\mu^2+\sigma^2)^{3/2}\sqrt{\lambda}}+\frac{0.3051}{\sqrt{n}}\right]=\frac{0.3051\beta^3}{(\mu^2+\sigma^2)^{3/2}\sqrt{\lambda}},$$

Q. E. D.

4 Convergence rate estimates in limit theorems for mixed compound Poisson distributions

4.1 Preliminaries

Let Λ_t be a positive random variable whose distribution depends on some parameter $t > 0$. The distribution function of Λ_t will be denoted $G_t(x)$: $G_t(x) = \mathsf{P}(\Lambda_t < x)$. By a mixed Poisson distribution with a structural distribution G_t we will mean the distribution of the random variable $N(t)$ which takes values $k = 0, 1, ...$ with probabilities

$$\mathsf{P}\big(N(t) = k\big) = \frac{1}{k!}\int_0^\infty e^{-\lambda}\lambda^k dG_t(\lambda), \qquad k = 0, 1, 2, ...$$

Some special examples of mixed Poisson distributions are very well-known. The most well-known and most widely used mixed Poisson distribution is, of course, the negative binomial distribution (since it was first used in the form of a mixed Poisson distribution in (Greenwood and Yule, 1920) to model the frequencies of accidents). This distribution is generated by the structural gamma-distribution. Other examples of mixed Poisson distributions are the Delaporte distribution with the shifted gamma-structural distribution (Delaporte, 1960), the Sichel distribution with the generalized inverse Gaussian structural distribution (Holla, 1967), (Sichel, 1971), Willmot, 1987), The generalized Waring distribution (Irwin, 1968), (Seal, 1978). The properties of mixed Poisson distributions are described in detail in (Grandell, 1997) and (Bening and Korolev, 2002).

Let $X_1, X_2, ...$ be independent identically distributed random variables. Assume that the random variables $N(t), X_1, X_2, ...$ are independent for each $t > 0$. Set

$$S(t) = X_1 + \ldots + X_{N(t)}$$

(for definiteness we assume that if $N(t) = 0$, then $S(t) = 0$). The random variable $S(t)$ will be called a mixed Poisson random sum and its distribution will be called compound mixed Poisson.

In what follows we will assume that the random variables $X_1, X_2, ...$ possess three first moments for which we will use the same notation as in Section 3 (see (13)). The asymptotic behavior of the distributions of mixed Poisson random sums $S(t)$ when $N(t)$ infinitely grows in some sense, is principally different depending on whether $\mu = 0$ or not.

The convergence in distribution and in probability will be respectively denoted by the symbols $\Longrightarrow$ and $\xrightarrow{P}$.

First consider the case $\mu = 0$. In this case the limit distributions for standardized mixed Poisson sums are scale mixtures of normal laws. Without loss of generality, unless otherwise indicated, we will assume that $\sigma^2 = 1$.

Theorem 3 (Korolev, 1996), (Bening and Korolev, 2002). *Assume that $\Lambda_t \xrightarrow{P} \infty$ as $t \to \infty$. Then, for a positive infinitely increasing function $d(t)$ there exists a distribution function $H(x)$ such that*

$$\mathsf{P}\left(\frac{S(t)}{\sqrt{d(t)}} < x\right) \Longrightarrow H(x) \qquad (t \to \infty)$$

if and only if there exists a distribution function $G(x)$ such that for the same function $d(t)$

$$G_t\big(xd(t)\big) \Longrightarrow G(x) \quad (t \to \infty) \tag{17}$$

and

$$H(x) = \int_0^\infty \Phi\big(x/\sqrt{y}\big)dG(y), \quad x \in \mathbb{R}.$$

Now consider the case $\mu \neq 0$. This case is important from the point of view of insurance applications. Recall that, in general, $\mathsf{D}X_1 = \sigma^2$. Assume that there exist numbers $\ell \in (0,\infty)$ and $s \in (0,\infty)$ such that

$$\mathsf{E}\Lambda_t \equiv \ell t, \quad \mathsf{D}\Lambda_t \equiv s^2 t, \quad t > 0. \tag{18}$$

Then it is easy to make sure that

$$\mathsf{E}S(t) = \mu\ell t, \quad \mathsf{D}S(t) = [\ell(\mu^2 + \sigma^2) + \mu^2 s^2]t.$$

In the book (Bening and Korolev, 2002) a general theorem presenting necessary and sufficient conditions for the convergence of compound mixed Poisson distributions was proved. The following theorem is a particular case of that result.

THEOREM 4 (Bening and Korolev, 2002). *Let $\mu \neq 0$. In addition to the conditions on the moments of the structural random variable Λ_t assume that $\Lambda_t \xrightarrow{P} \infty$ as $t \to \infty$. Then, as $t \to \infty$, compound mixed Poisson distributions converge to the distribution of some random variable Z, that is,*

$$\frac{S(t) - \mu\ell t}{\sqrt{[\ell(\mu^2 + \sigma^2) + \mu^2 s^2]t}} \Longrightarrow Z,$$

if and only if there exists a random variable V such that

$$\frac{\Lambda_t - \ell t}{s\sqrt{t}} \Longrightarrow V.$$

Furthermore,

$$\mathsf{P}(Z < x) = \mathsf{E}\Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2 + \sigma^2)\ell}} - \frac{\mu s V}{\sqrt{(\sigma^2 + \mu^2)\ell}}\right), \quad x \in \mathbb{R}.$$

It is easy to see that the limit random variable Z admits the representation

$$Z \stackrel{d}{=} \left[1 + \frac{\mu^2 s^2}{(\mu^2 + \sigma^2)\ell}\right]^{-1/2} \cdot X + \frac{\mu s}{\sqrt{(\mu^2 + \sigma^2)\ell + \mu^2 s^2}} \cdot V,$$

where X is a random variable with the standard normal distribution independent of V.

The basic distinctions of the case $\mu \neq 0$ from the case of compound mixed Poisson distributions with zero expectations considered above are, first, the necessity of non-trivial centering and different normalization required for the existence of non-trivial limit laws and, second, the shape of the limit law which in this case has the form of a location mixture of normal laws.

4.2 Convergence rate estimates in limit theorems for mixed compound Poisson distributions with zero mean

It is easily seen that the distribution of the mixed Poisson random sum $S(t)$ can be represented as

$$\mathsf{P}(S(t) < x) = \int_0^\infty \mathsf{P}\Big(\sum_{j=1}^{N_\lambda} X_j < x\Big) dG_t(\lambda), \quad x \in \mathbb{R}. \tag{19}$$

Recall that here we assume that

$$\mathsf{E}X_1 = 0, \quad \mathsf{E}X_1^2 = 1, \quad \beta^3 = \mathsf{E}|X_1|^3 < \infty. \tag{20}$$

Let $d(t)$, $t > 0$, be a positive infinitely increasing function. In this section we will present some estimates of the rate of convergence in theorem 3.

For $\lambda > 0$ denote

$$\rho(\lambda) = \sup_x \bigg|\mathsf{P}\Big(\frac{1}{\sqrt{\lambda}}\sum_{j=1}^{N_\lambda} X_j < x\Big) - \Phi(x)\bigg|.$$

Let $G(x)$ be a distribution function such that $G(0) = 0$. If condition (17) holds, then, according to theorem 3, compound mixed Poisson distribution of the mixed Poisson sum $S(t)$ normalized by the square root of the function $d(t)$ converges to the scale mixture of normal laws in which $G(x)$ is the mixing distribution. Denote

$$\Delta_t = \sup_x \bigg|\mathsf{P}\Big(\frac{S(t)}{\sqrt{d(t)}} < x\Big) - \int_0^\infty \Phi\Big(\frac{x}{\sqrt{\lambda}}\Big) dG(\lambda)\bigg|, \quad \delta_t = \sup_x \big|G_t\big(d(t)x\big) - G(x)\big|.$$

THEOREM 5. *Assume that conditions* (20) *hold. Then for any* $t > 0$ *we have the estimate*

$$\Delta_t \leqslant 0.3051 \cdot \beta^3 \mathsf{E}[\Lambda_t]^{-1/2} + 0.5 \cdot \delta_t.$$

P r o o f. This statement was first proved in the paper (Gavrilenko and Korolev, 2006) with a slightly worse constant (also see (Korolev, Bening and Shorgin, 2007). Here we present a modified version of the proof. By virtue of representation (19) we have

$$\Delta_t = \sup_x \bigg|\int_0^\infty \mathsf{P}\Big(\sum_{j=1}^{N_\lambda} X_j < x\sqrt{d(t)}\Big) dG_t(\lambda) - \int_0^\infty \Phi\Big(\frac{x}{\sqrt{\lambda}}\Big) dG(\lambda)\bigg| =$$

$$= \sup_x \bigg|\int_0^\infty \mathsf{P}\Big(\frac{1}{\sqrt{\lambda}}\sum_{j=1}^{N_\lambda} X_j < x\frac{\sqrt{d(t)}}{\sqrt{\lambda}}\Big) dG_t(\lambda) - \int_0^\infty \Phi\Big(\frac{x}{\sqrt{\lambda}}\Big) dG(\lambda)\bigg| =$$

$$= \sup_x \bigg|\int_0^\infty \mathsf{P}\Big(\frac{1}{\sqrt{\lambda d(t)}}\sum_{j=1}^{N_{\lambda d(t)}} X_j < \frac{x}{\sqrt{\lambda}}\Big) dG_t\big(\lambda d(t)\big)\Big) - \int_0^\infty \Phi\Big(\frac{x}{\sqrt{\lambda}}\Big) dG(\lambda)\bigg| \leqslant$$

$$\leqslant \sup_x \left| \int\limits_0^\infty \left[\mathsf{P}\left(\frac{1}{\sqrt{\lambda d(t)}} \sum_{j=1}^{N_{\lambda d(t)}} X_j < \frac{x}{\sqrt{\lambda}} \right) - \Phi\left(\frac{x}{\sqrt{\lambda}} \right) \right] dG_t\big(\lambda d(t)\big) \right| +$$

$$+ \sup_x \left| \int\limits_0^\infty \Phi\left(\frac{x}{\sqrt{\lambda}} \right) d\big[G_t\big(\lambda d(t)\big) - G(\lambda)\big] \right|.$$

Continuing this chain of relations with the use of integration by parts and theorem 2 we further obtain

$$\Delta_t \leqslant \int\limits_0^\infty \sup_x \left| \mathsf{P}\left(\frac{1}{\sqrt{\lambda d(t)}} \sum_{j=1}^{N_{\lambda d(t)}} X_j < x \right) - \Phi(x) \right| dG_t\big(\lambda d(t)\big) +$$

$$+ \sup_x \left| \int\limits_0^\infty \big[G_t\big(\lambda d(t)\big) - G(\lambda)\big] d_\lambda \Phi\left(\frac{x}{\sqrt{\lambda}} \right) \right| \leqslant$$

$$\leqslant \int\limits_0^\infty \rho(\lambda) dG_t(\lambda) + \sup_\lambda \big|G_t\big(\lambda d(t)\big) - G(\lambda)\big| \cdot \sup_x \left| \int\limits_0^\infty d_\lambda \Phi\left(\frac{x}{\sqrt{\lambda}} \right) \right| \leqslant$$

$$\leqslant 0.3051 \cdot \beta^3 \int\limits_0^\infty \frac{1}{\sqrt{\lambda}} dG_t(\lambda) + 0.5 \cdot \sup_\lambda \big|G_t\big(\lambda d(t)\big) - G(\lambda)\big| =$$

$$= 0.3051 \cdot \beta^3 \mathsf{E}[\Lambda_t]^{-1/2} + 0.5 \cdot \delta_t,$$

Q. E. D.

As an example of applications of theorem 5 consider the case where for each $t > 0$ the random variable Λ_t has the gamma-distribution. This case is very important in financial applications for the asymptotic validation of such popular models of the evolution of financial indexes as variance-gamma Lévy processes (VG-processes) (Madan and Seneta, 1990) or two-sided gamma-processes (Carr, Madan and Chang, 1998).

As is well known, the density of the gamma-distribution with shape parameter $r > 0$ and scale parameter $\sigma > 0$ has the form

$$g_{r,\sigma}(x) = \frac{\sigma^r}{\Gamma(r)} e^{-\sigma x} x^{r-1}, \quad x > 0.$$

Thus, the mixed Poisson distribution with the mixing gamma-distribution has the characteristic function

$$\psi(t) = \int\limits_0^\infty \exp\{y(e^{it} - 1)\} \frac{\sigma^r}{\Gamma(r)} e^{-\sigma y} y^{r-1} dy =$$

$$= \frac{\sigma^r}{\Gamma(r)} \int\limits_0^\infty \exp\left\{ -\sigma y \left(1 + \frac{1 - e^{it}}{\sigma} \right) \right\} y^{r-1} dy = \left(1 + \frac{1 - e^{it}}{\sigma} \right)^{-r}.$$

By the re-parametrization

$$\sigma = \frac{p}{1-p} \quad \Big(p = \frac{\sigma}{1+\sigma}\Big), \quad p \in (0,1),$$

we finally obtain

$$\psi(t) = \Big(\frac{p}{1-(1-p)e^{it}}\Big)^r, \quad t \in \mathbb{R},$$

which coincides with the characteristic function of the negative binomial distribution with parameters $r > 0$ and $p \in (0,1)$. So, in the case under consideration for each $t > 0$ the random variable $N(t)$ has the negative binomial distribution with parameters $r > 0$ and $p \in (0,1)$:

$$\mathsf{P}\big(N(t) = n\big) = C^n_{r+n-1} p^r (1-p)^n, \qquad n = 0, 1, 2, \ldots. \tag{21}$$

Here $r > 0$ and $p \in (0,1)$ are the parameters and for non-integer r the quantity C^n_{r+n-1} is defined as

$$C^n_{r+n-1} \;=\; \frac{\Gamma(r+n)}{n! \cdot \Gamma(r)}.$$

In particular, with $r = 1$, relation (21) determines the geometric distribution.

The gamma-distribution function with scale parameter σ ana shape parameter r will be denoted $G_{r,\sigma}(x)$. It is easy to see that

$$G_{r,\sigma}(x) \equiv G_{r,1}(\sigma x). \tag{22}$$

The random variable with the distribution function $G_{r,\sigma}(x)$ will be denoted $U(r,\sigma)$. It is well known that

$$\mathsf{E}U(r,\sigma) = \frac{r}{\sigma}.$$

Fix the parameter r and take $U(r,\sigma)$ as the random variable Λ_t assuming that $t = \sigma^{-1}$:

$$\Lambda_t = U(r, t^{-1}).$$

As a function $d(t)$ take

$$d(t) \equiv \mathsf{E}\Lambda_t = \mathsf{E}U(r, t^{-1}).$$

Obviously, we have

$$\mathsf{E}U(r, t^{-1}) = rt.$$

Then with the account of (22) we have

$$G_t\big(xd(t)\big) = \mathsf{P}(U(r,t^{-1}) < xrt) = \mathsf{P}(U(r,1) < xr) = \mathsf{P}(U(r,r) < x) = G_{r,r}(x).$$

Note that the distribution function on the right-hand side of the latter relation does not depend on t. Therefore the choice of $d(t)$ specified above trivially guarantees the validity of condition (17) of theorem 3. Moreover, in this case $\delta_t = 0$ for all $t > 0$.

Now calculate $\mathsf{E}[\Lambda_t]^{-1/2}$ under the condition

$$r > \frac{1}{2}. \tag{23}$$

We have

$$\mathsf{E}[\Lambda_t]^{-1/2} = \mathsf{E}[U(r,t^{-1})]^{-1/2} = \int\limits_0^\infty \frac{e^{-x/t}x^{r-3/2}}{t^r\Gamma(r)}dx = \frac{\Gamma(r-\frac{1}{2})}{\Gamma(r)\sqrt{t}}.$$

Thus we obtain the following statement which is actually a particular case of theorem 5.

COROLLARY 2. *Let the random variable Λ_t have the gamma-distribution with shape parameter $r > 0$ and scale parameter $\sigma = 1/t$, $t > 0$. Assume that conditions* (20) *and* (23) *hold. Then for each $t > 0$ we have*

$$\sup_x \left|\mathsf{P}(S(t) < x\sqrt{rt}) - \int\limits_0^\infty \Phi\left(\frac{x}{\sqrt{y}}\right)dG_{r,r}(y)\right| \leqslant 0.3051\frac{\Gamma(r-\frac{1}{2})}{\Gamma(r)}\cdot\frac{\beta^3}{\sqrt{t}}.$$

If $r = 1$, then the random variable

$$N(t) = N_1\big(U(1,t^{-1})\big), \quad t \geqslant 0,$$

has the geometric distribution with parameter $p = t^{-1}$. As this is so, the limit (as $t \to \infty$) distribution function of the standardized geometric sum $S(t)$ is the Laplace distribution with the density

$$l(x) = \frac{1}{\sqrt{2}}e^{-\sqrt{2}|x|}, \qquad x \in \mathbb{R}.$$

The distribution function corresponding to the density $l(x)$ will be denoted $L(x)$,

$$L(x) = \begin{cases} \frac{1}{2}e^{\sqrt{2}x}, & \text{if } x < 0, \\ 1 - \frac{1}{2}e^{-\sqrt{2}x}, & \text{if } x \geqslant 0. \end{cases}$$

COROLLARY 3. *Let the random variable Λ_t have the exponential distribution with parameter $\sigma = 1/t$, $t > 0$. Assume that conditions* (20) *hold. Then for each $t > 0$*

$$\sup_x |\mathsf{P}(S(t) < x\sqrt{t}) - L(x)| \leqslant 0.5408\cdot\frac{\beta_3}{\sqrt{t}}.$$

5 Convergence rate estimates in limit theorems for mixed compound Poisson distributions with non-zero mean

Here we will present some estimates of the rate of convergence in theorem 4.

5.1 The case of structural random variables with finite variance

Under assumptions (18) denote

$$F_t(x) = \mathsf{P}\left(\frac{S(t) - \mu\ell t}{\sqrt{[\ell(\mu^2+\sigma^2)+\mu^2 s^2]t}} < x\right),$$

$$\rho_t = \sup_x \left| F_t(x) - \mathsf{E}\Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s V}{\sqrt{(\sigma^2+\mu^2)\ell}} \right) \right|,$$

$$G^*(v) = \mathsf{P}\,(V < v)\,, \quad \widetilde{\delta}_t = \sup_v \Big| G_t\big(vs\sqrt{t} + \ell t\big) - G^*(v) \Big|.$$

THEOREM 6. *Let* $\mathsf{E}X_1 = \mu \neq 0$, $\mathsf{D}X_1 = \sigma^2$, $\mathsf{E}|X_1|^3 = \beta^3 < \infty$, $\mathsf{E}\,|V| < \infty$. *Then for any* $t > 0$ *we have*

$$\rho_t \le \widetilde{\delta}_t + \frac{1}{\sqrt{t}} \cdot \inf_{\epsilon \in (0,1)} \left\{ \frac{0.3051\beta^3}{(\mu^2+\sigma^2)^{3/2}\sqrt{(1-\epsilon)\ell}} + \frac{s}{\ell}\left(\frac{\mathsf{E}|V|}{\epsilon} + Q(\epsilon) \right) \Bigg| \right\},$$

where

$$Q(\epsilon) = \max\left\{ \frac{1}{\epsilon},\ \frac{\sqrt{1+\epsilon}}{\big(1+\sqrt{1-\epsilon}\big)\sqrt{2\pi e(1-\epsilon)}} \right\}.$$

P r o o f. A similar statement with slightly worse constants was first proved in the paper (Artyukhov and Korolev, 2008). Here we present a modified version of the proof. As above, let N_λ be a random variable with the Poisson distribution with parameter $\lambda > 0$ independent of the sequence $X_1, X_2, \ldots$ Then we can write

$$\rho_t = \sup_x \left| \mathsf{P}\left(\frac{S(t) - \mu\ell t}{\sqrt{[\ell(\mu^2+\sigma^2) + \mu^2 s^2]t}} < x \right) - \mathsf{E}\,\Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s V}{\sqrt{(\mu^2+\sigma^2)\ell}} \right) \right| =$$

$$= \sup_x \left| \int_0^\infty \mathsf{P}\left(\frac{S_{N_\lambda} - \mu\ell t}{\sqrt{[\ell(\mu^2+\sigma^2) + \mu^2 s^2]t}} < x \right) dG_t(\lambda) - \right.$$

$$\left. - \mathsf{E}\,\Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s V}{\sqrt{(\mu^2+\sigma^2)\ell}} \right) \right|.$$

Fix an arbitrary $\epsilon \in (0,1)$. Then

$$\rho_t = \sup_x \left| \int\limits_{\lambda < (1-\epsilon)\ell t} \mathsf{P}\left(\frac{S_{N_\lambda} - \mu\ell t}{\sqrt{[\ell(\mu^2+\sigma^2) + \mu^2 s^2]t}} < x \right) dG_t(\lambda) + \right.$$

$$+ \int\limits_{\lambda > (1+\epsilon)\ell t} \mathsf{P}\left(\frac{S_{N_\lambda} - \mu\ell t}{\sqrt{[\ell(\mu^2+\sigma^2) + \mu^2 s^2]t}} < x \right) dG_t(\lambda) +$$

$$+ \int\limits_{(1-\epsilon)\ell t \leqslant \lambda \leqslant (1+\epsilon)\ell t} \mathsf{P}\left(\frac{S_{N_\lambda} - \mu\ell t}{\sqrt{[\ell(\mu^2+\sigma^2) + \mu^2 s^2]t}} < x \right) dG_t(\lambda) -$$

$$\left. - \mathsf{E}\,\Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s V}{\sqrt{(\mu^2+\sigma^2)\ell}} \right) \right| \leqslant$$

$$\leqslant \sup_x \left| \int\limits_{\lambda<(1-\epsilon)\ell t} \mathsf{P}\left(\frac{S_{N_\lambda} - \mu\ell t}{\sqrt{[\ell(\mu^2+\sigma^2)+\mu^2 s^2]t}} < x \right) dG_t(\lambda) + \right.$$

$$\left. + \int\limits_{\lambda>(1+\epsilon)\ell t} \mathsf{P}\left(\frac{S_{N_\lambda} - \mu\ell t}{\sqrt{[\ell(\mu^2+\sigma^2)+\mu^2 s^2]t}} < x \right) dG_t(\lambda) \right| +$$

$$+ \sup_x \left| \int\limits_{(1-\epsilon)\ell t \leqslant \lambda \leqslant (1+\epsilon)\ell t} \mathsf{P}\left(\frac{S_{N_\lambda} - \mu\ell t}{\sqrt{[\ell(\mu^2+\sigma^2)+\mu^2 s^2]t}} < x \right) dG_t(\lambda) - \right.$$

$$\left. - \mathsf{E}\,\Phi\left(x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s V}{\sqrt{(\mu^2+\sigma^2)\ell}} \right) \right| \leqslant$$

$$\leqslant \mathsf{P}\left(\left| \frac{\Lambda_t}{\ell t} - 1 \right| > \epsilon \right) +$$

$$+ \sup_x \left| \int\limits_{(1-\epsilon)\ell t \leqslant \lambda \leqslant (1+\epsilon)\ell t} \mathsf{P}\left(\frac{S_{N_\lambda} - \mu\ell t}{\sqrt{[\ell(\mu^2+\sigma^2)+\mu^2 s^2]t}} < x \right) dG_t(\lambda) - \right.$$

$$\left. - \mathsf{E}\,\Phi\left(x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s V}{\sqrt{(\mu^2+\sigma^2)\ell}} \right) \right|. \qquad (24)$$

Further,

$$\sup_x \left| \int\limits_{(1-\epsilon)\ell t \leqslant \lambda \leqslant (1+\epsilon)\ell t} \mathsf{P}\left(\frac{S_{N_\lambda} - \mu\ell t}{\sqrt{[\ell(\mu^2+\sigma^2)+\mu^2 s^2]t}} < x \right) dG_t(\lambda) - \right.$$

$$\left. - \mathsf{E}\,\Phi\left(x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s V}{\sqrt{(\mu^2+\sigma^2)\ell}} \right) \right| \leqslant$$

$$\leqslant \sup_x \int\limits_{(1-\epsilon)\ell t \leqslant \lambda \leqslant (1+\epsilon)\ell t} \left| \mathsf{P}\left(\frac{S_{N_\lambda} - \mu\lambda}{\sqrt{\lambda(\mu^2+\sigma^2)}} < \right.\right.$$

$$\left. < \sqrt{\frac{\ell t}{\lambda}}\left(x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s}{\sqrt{(\mu^2+\sigma^2)\ell}} \cdot \frac{\lambda-\ell t}{s\sqrt{t}} \right) \right) -$$

$$\left. - \Phi\left(\sqrt{\frac{\ell t}{\lambda}}\left(x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s}{\sqrt{(\mu^2+\sigma^2)\ell}} \cdot \frac{\lambda-\ell t}{s\sqrt{t}} \right) \right) \right| dG_t(\lambda) +$$

$$+ \sup_x \left| \int\limits_{(1-\epsilon)\ell t \leqslant \lambda \leqslant (1+\epsilon)\ell t} \Phi\left(\sqrt{\frac{\ell t}{\lambda}}\left(x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s}{\sqrt{(\mu^2+\sigma^2)\ell}} \cdot \frac{\lambda-\ell t}{s\sqrt{t}} \right) \right) dG_t(\lambda) - \right.$$

$$\left. - \mathsf{E}\,\Phi\left(x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s V}{\sqrt{(\mu^2+\sigma^2)\ell}} \right) \right| \equiv I_1 + I_2.$$

Consider I_1. Denote

$$y = \sqrt{\frac{\ell t}{\lambda}}\left(x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s}{\sqrt{(\mu^2+\sigma^2)\ell}}\cdot\frac{\lambda-\ell t}{s\sqrt{t}}\right).$$

Then I_1 can be rewritten in the form

$$I_1 = \sup_y \int\limits_{(1-\epsilon)\ell t\leqslant\lambda\leqslant(1+\epsilon)\ell t} \left|\mathsf{P}\left(\frac{S_{N_\lambda}-\mu\lambda}{\sqrt{\lambda(\mu^2+\sigma^2)}} < y\right) - \Phi(y)\right| dG_t(\lambda) \leqslant$$

$$\leqslant \int\limits_{(1-\epsilon)\ell t\leqslant\lambda\leqslant(1+\epsilon)\ell t} \sup_y \left|\mathsf{P}\left(\frac{S_{N_\lambda}-\mu\lambda}{\sqrt{\lambda(\mu^2+\sigma^2)}} < y\right) - \Phi(y)\right| dG_t(\lambda).$$

To estimate the integrand on the right-hand side of the latter inequality we use theorem 2 and obtain

$$I_1 \leqslant \frac{0.3051\beta^3}{(\mu^2+\sigma^2)^{3/2}} \int\limits_{\lambda\geqslant(1-\epsilon)\ell t} \frac{1}{\sqrt{\lambda}}\,dG_t(\lambda) \leqslant \frac{0.3051\beta^3}{(\mu^2+\sigma^2)^{3/2}\sqrt{(1-\epsilon)\ell t}}. \tag{25}$$

Consider I_2. We have

$$I_2 \leqslant \sup_x \int\limits_{(1-\epsilon)\ell t\leqslant\lambda\leqslant(1+\epsilon)\ell t} \left|\Phi\left(\sqrt{\frac{\ell t}{\lambda}}\left(x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s}{\sqrt{(\mu^2+\sigma^2)\ell}}\cdot\frac{\lambda-\ell t}{s\sqrt{t}}\right)\right) - \right.$$

$$\left. -\Phi\left(x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s}{\sqrt{(\mu^2+\sigma^2)\ell}}\cdot\frac{\lambda-\ell t}{s\sqrt{t}}\right)\right| dG_t(\lambda) +$$

$$+\sup_x \left|\int\limits_{(1-\epsilon)\ell t\leqslant\lambda\leqslant(1+\epsilon)\ell t} \Phi\left(x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s}{\sqrt{(\mu^2+\sigma^2)\ell}}\cdot\frac{\lambda-\ell t}{s\sqrt{t}}\right) dG_t(\lambda) - \right.$$

$$\left. -\mathsf{E}\,\Phi\left(x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s V}{\sqrt{(\mu^2+\sigma^2)\ell}}\right)\right| \equiv I_{21} + I_{22}.$$

Denote

$$z = x\sqrt{1+\frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s}{\sqrt{(\mu^2+\sigma^2)\ell}}\cdot\frac{\lambda-\ell t}{s\sqrt{t}}.$$

Then

$$I_{21} \leqslant \int\limits_{(1-\epsilon)\ell t\leqslant\lambda\leqslant(1+\epsilon)\ell t} \sup_z \left|\Phi\left(z\sqrt{\frac{\ell t}{\lambda}}\right) - \Phi(z)\right| dG_t(\lambda). \tag{26}$$

Consider the integrand in (26). By the Lagrange formula we have

$$\left|\Phi\left(z\sqrt{\frac{\ell t}{\lambda}}\right) - \Phi(z)\right| = |z|\cdot\left|\sqrt{\frac{\ell t}{\lambda}} - 1\right|\varphi\left(\theta z + (1-\theta)z\sqrt{\frac{\ell t}{\lambda}}\right) \tag{27}$$

for some $\theta \in [0,1]$ where

$$\varphi(x) = \Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

is the standard normal density. The function $\varphi(x) = \varphi(|x|)$ monotonically decreases as $|x|$ increases. Therefore the function φ on the right-hand side of (27) attains its maximum value in $\theta \in [0,1]$ at that value of its argument, whose absolute value is minimum. But the argument of the function φ on the right-hand side of (27) is itself a linear function of θ. Therefore, the minimum absolute value of this argument is attained either at $\theta = 0$ or at $\theta = 1$. But at $\theta = 1$ we have

$$\theta\left(1 - \sqrt{\frac{\ell t}{\lambda}}\right) + \sqrt{\frac{\ell t}{\lambda}} = 1,$$

while at $\theta = 0$ we have

$$\theta\left(1 - \sqrt{\frac{\ell t}{\lambda}}\right) + \sqrt{\frac{\ell t}{\lambda}} = \sqrt{\frac{\ell t}{\lambda}}.$$

In the definition of I_{21} λ satisfies the inequality $\lambda \leqslant (1+\epsilon)\ell t$. Therefore,

$$\sqrt{\frac{\ell t}{\lambda}} \geqslant \frac{1}{\sqrt{1+\epsilon}}.$$

Hence,

$$\theta\left(1 - \sqrt{\frac{\ell t}{\lambda}}\right) + \sqrt{\frac{\ell t}{\lambda}} \geqslant \min\left\{1, \frac{1}{\sqrt{1+\epsilon}}\right\} = \frac{1}{\sqrt{1+\epsilon}}.$$

Therefore in I_{21} we have (see (27))

$$\sup_z \left|\Phi\left(z\sqrt{\frac{\ell t}{\lambda}}\right) - \Phi(z)\right| \leqslant \left|\sqrt{\frac{\ell t}{\lambda}} - 1\right| \cdot \sup_z |z|\varphi\left(\frac{z}{\sqrt{1+\epsilon}}\right). \tag{28}$$

Furthermore,

$$\left(z\varphi\left(\frac{z}{\sqrt{1+\epsilon}}\right)\right)' = \varphi\left(\frac{z}{\sqrt{1+\epsilon}}\right)\left(1 - \frac{z^2}{1+\epsilon}\right).$$

Therefore the supremum in (28) is attained at $z = \pm\sqrt{1+\epsilon}$ and equals

$$\sup_z |z|\varphi\left(\frac{z}{\sqrt{1+\epsilon}}\right) = \sqrt{\frac{1+\epsilon}{2\pi e}}.$$

Thus,

$$I_{21} \leqslant \sqrt{\frac{1+\epsilon}{2\pi e}} \int\limits_{(1-\epsilon)\ell t \leqslant \lambda \leqslant (1+\epsilon)\ell t} \left|\sqrt{\frac{\ell t}{\lambda}} - 1\right| dG_t(\lambda) =$$

$$= \sqrt{\frac{1+\epsilon}{2\pi e}} \int\limits_{(1-\epsilon)\ell t \leqslant \lambda \leqslant (1+\epsilon)\ell t} \left|\frac{\sqrt{\ell t} - \sqrt{\lambda}}{\sqrt{\lambda}}\right| \cdot \left|\frac{\sqrt{\ell t} + \sqrt{\lambda}}{\sqrt{\ell t} + \sqrt{\lambda}}\right| dG_t(\lambda) =$$

$$= \sqrt{\frac{1+\epsilon}{2\pi e}} \int\limits_{(1-\epsilon)\ell t \leqslant \lambda \leqslant (1+\epsilon)\ell t} \frac{|\lambda - \ell t|}{\sqrt{\lambda}\big(\sqrt{\ell t} + \sqrt{\lambda}\big)}\, dG_t(\lambda) \leqslant$$

$$\leqslant \sqrt{\frac{1+\epsilon}{2\pi e}} \cdot \frac{1}{\sqrt{1-\epsilon}\big(1+\sqrt{1-\epsilon}\big)} \int\limits_{(1-\epsilon)\ell t \leqslant \lambda \leqslant (1+\epsilon)\ell t} \left|\frac{\lambda}{\ell t} - 1\right| dG_t(\lambda) =$$

$$= \sqrt{\frac{1+\epsilon}{2\pi e(1-\epsilon)}} \cdot \frac{1}{1+\sqrt{1-\epsilon}} \cdot \mathsf{E}\left|\frac{\Lambda_t}{\ell t} - 1\right| \mathbf{1}\left(\left|\frac{\Lambda_t}{\ell t} - 1\right| \leqslant \epsilon\right). \tag{29}$$

Here the symbol $\mathbf{1}(A)$ denotes the indicator of a set A.

Consider I_{22}. We have

$$\int\limits_{\lambda:\, |\frac{\lambda}{\ell t} - 1| \leqslant \epsilon} \Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s}{\sqrt{(\mu^2+\sigma^2)\ell}} \cdot \frac{\lambda - \ell t}{s\sqrt{t}}\right) dG_t(\lambda) =$$

$$= \int\limits_{\lambda:\, |\frac{\lambda - \ell t}{s\sqrt{t}}| \leqslant \epsilon \ell\sqrt{t}/s} \Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s}{\sqrt{(\mu^2+\sigma^2)\ell}} \cdot \frac{\lambda - \ell t}{s\sqrt{t}}\right) dG_t(\lambda) =$$

$$= \int\limits_{|v| \leqslant \epsilon \ell\sqrt{t}/s} \Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s v}{\sqrt{(\mu^2+\sigma^2)\ell}}\right) dG_t(vs\sqrt{t} + \ell t).$$

Therefore,

$$I_{22} = \sup_x \left| \int\limits_{|v| \leqslant \epsilon \ell\sqrt{t}/s} \Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s v}{\sqrt{(\mu^2+\sigma^2)\ell}}\right) dG_t(vs\sqrt{t} + \ell t) - \right.$$

$$\left. - \int\limits_{-\infty}^{\infty} \Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s v}{\sqrt{(\mu^2+\sigma^2)\ell}}\right) dG^*(v) \right| \leqslant$$

$$\leqslant \sup_x \left| \int\limits_{|v| \leqslant \epsilon \ell\sqrt{t}/s} \Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s v}{\sqrt{(\mu^2+\sigma^2)\ell}}\right) \big[dG_t(vs\sqrt{t} + \ell t) - G^*(v)\big] \right| +$$

$$+ \sup_x \left| \int\limits_{|v| > \epsilon \ell\sqrt{t}/s} \Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s v}{\sqrt{(\mu^2+\sigma^2)\ell}}\right) dG^*(v) \right| \equiv I_{221} + I_{222}.$$

By integration by parts we obtain

$$I_{221} \leqslant \sup_x \left| \int\limits_{|v| \leqslant \epsilon \ell\sqrt{t}/s} \big[G_t(vs\sqrt{t} + \ell t) - G^*(v)\big]\, d_v \Phi\left(x\sqrt{1 + \frac{\mu^2 s^2}{(\mu^2+\sigma^2)\ell}} - \frac{\mu s v}{\sqrt{(\mu^2+\sigma^2)\ell}}\right) \right| \leqslant$$

$$\leqslant \sup_v \left| G_t(vs\sqrt{t}+\ell t) - G^*(v) \right| \equiv \widetilde{\delta}_t. \tag{30}$$

Note that in (24) we can apply the Markov inequality and obtain the estimate

$$\mathsf{P}\left(\left|\frac{\Lambda_t}{\ell t}-1\right| > \epsilon\right) \leqslant \frac{1}{\epsilon}\cdot \mathsf{E}\left|\frac{\Lambda_t}{\ell t}-1\right| \mathbf{1}\left(\left|\frac{\Lambda_t}{\ell t}-1\right| > \epsilon\right). \tag{31}$$

Further, again applying the Markov inequality we can make sure that

$$I_{222} \leqslant \mathsf{P}\big(|V| > \epsilon\ell\sqrt{t}/s\big) \leqslant \frac{s\mathsf{E}|V|}{\epsilon\ell\sqrt{t}}. \tag{32}$$

Now unifying (24), (25), (26), (29), (30), (31) and (32) we finally obtain that for any $\epsilon \in (0,1)$ there holds the inequality

$$\rho_t \leqslant \widetilde{\delta}_t + \frac{1}{\sqrt{t}}\left(\frac{0.3051\beta^3}{(\mu^2+\sigma^2)^{3/2}\sqrt{(1-\epsilon)\ell}} + \frac{s}{\epsilon\ell}\cdot \mathsf{E}|V|\right) +$$

$$+\mathsf{E}\left|\frac{\Lambda_t}{\ell t}-1\right| \cdot \max\left\{\frac{1}{\epsilon}, \sqrt{\frac{1+\epsilon}{2\pi e(1-\epsilon)}}\cdot\frac{1}{1+\sqrt{1-\epsilon}}\right\},$$

whence we obviously obtain the statement of the theorem since the Lyapunov inequality obviously implies that for each $t>0$

$$\mathsf{E}\left|\frac{\Lambda_t}{\ell t}-1\right| = \frac{s}{\ell\sqrt{t}}\,\mathsf{E}\left|\frac{\Lambda_t-\ell t}{s\sqrt{t}}\right| \leqslant \frac{s}{\ell\sqrt{t}}\sqrt{\mathsf{D}\left(\frac{\Lambda_t-\ell t}{s\sqrt{t}}\right)} = \frac{s}{\ell\sqrt{t}}.$$

The theorem is proved.

If we additionally assume that the family of random variables

$$\left\{\left|\frac{\Lambda_t-\ell t}{s\sqrt{t}}\right|\right\}_{t>0}$$

is uniformly integrable, then by the Lyapunov inequality we obtain the inequality

$$\mathsf{E}|V| = \lim_{t\to\infty}\mathsf{E}\left|\frac{\Lambda_t-\ell t}{s\sqrt{t}}\right| \leqslant \lim_{t\to\infty}\sqrt{\mathsf{D}\left(\frac{\Lambda_t-\ell t}{s\sqrt{t}}\right)} = 1. \tag{33}$$

Hence, from theorem 6 we obtain the following result.

Corollary 4. *In addition to the conditions of theorem* 6, *let* (33) *hold. Then for any* $t>0$

$$\rho_t \leq \widetilde{\delta}_t + \frac{1}{\sqrt{t}}\cdot\inf_{\epsilon\in(0,1)}\left\{\frac{0.3051\beta^3}{(\mu^2+\sigma^2)^{3/2}\sqrt{(1-\epsilon)\ell}} + \frac{s}{\ell}\left(\frac{1}{\epsilon}+Q(\epsilon)\right)\right\},$$

where

$$Q(\epsilon) = \max\left\{\frac{1}{\epsilon}, \frac{\sqrt{1+\epsilon}}{\big(1+\sqrt{1-\epsilon}\big)\sqrt{2\pi e(1-\epsilon)}}\right\}.$$

5.2 The case of structural random variables with infinite variance

Assumption (18) which guarantees the existence of the variance of the structural random variable Λ_t is not crucial. An analog of theorem 6 can be proved for the case where only the existence of the mathematical expectation of Λ_t is assumed. Namely, the following theorem holds.

THEOREM 7. *Let $\mu \neq 0$. Assume that $\mathsf{E}\Lambda_t \equiv t$ and $\Lambda_t \xrightarrow{P} \infty$ as $t \to \infty$. Then, as $t \to \infty$, the distributions of normalized mixed Poisson random sums converge to the distribution of some random variable Z, that is,*

$$\frac{S(t) - \mu}{\sqrt{t}} \Longrightarrow Z,$$

if and only if there exists a random variable V such that

$$\frac{\Lambda_t - t}{\sqrt{t}} \Longrightarrow V.$$

Moreover,

$$\mathsf{P}(Z < x) = \mathsf{E}\Phi\Big(\frac{x - \mu V}{\sqrt{\sigma^2 + \mu^2}}\Big), \quad x \in \mathbb{R}. \tag{34}$$

Relation (34) means that in theorem 8

$$Z \stackrel{d}{=} \sqrt{\mu^2 + \sigma^2} \cdot X + \mu V$$

where the random variables X and V are independent and X has the standard normal distribution.

By analogy with the notation introduced above, denote

$$\widetilde{\rho}_t = \sup_x \left| \mathsf{P}\left(\frac{S(t) - \mu t}{\sqrt{t}} < x\right) - \mathsf{E}\Phi\Big(\frac{x - \mu V}{\sqrt{\sigma^2 + \mu^2}}\Big) \right|,$$

$$\widehat{\delta}_t = \sup_v \Big| G_t(v\sqrt{t} + t)) - G^*(v) \Big|.$$

THEOREM 8. *Assume that $\beta^3 < \infty$, $\mathsf{E}\Lambda_t \equiv t$, $t > 0$, and $\mathsf{E}\,|V| < \infty$. Then*

$$\widetilde{\rho}_t \leq \widehat{\delta}_t + \frac{1}{\sqrt{t}} \cdot \inf_{\epsilon \in (0,1)} \left\{ \frac{0.3051\beta^3}{(\mu^2 + \sigma^2)^{3/2}\sqrt{1 - \epsilon}} + \frac{\mathsf{E}|V|}{\epsilon} + Q(\epsilon)\mathsf{E}\left|\frac{\Lambda_t - t}{\sqrt{t}}\right| \right\},$$

where

$$Q(\epsilon) = \max\left\{\frac{1}{\epsilon}, \frac{\sqrt{1 + \epsilon}}{\big(1 + \sqrt{1 - \epsilon}\big)\sqrt{2\pi e(1 - \epsilon)}}\right\}.$$

The proof of theorem 8 differs from the proof of theorem 6 only in notation.

As an example of the situation in which theorems 7 and 8 are valid, but theorems 4 and 6 are not, consider the case where

$$\Lambda_t = \max\{0, \sqrt{t}V + t\} + \frac{1}{2t^{\alpha/2}}\Big(\frac{2\alpha + 1}{\alpha}\sqrt{t} - 1\Big),$$

with $2 < \alpha < 3$ and V being the random variable with the density

$$p(x) = \frac{\alpha+1}{2(|x|+1)^{\alpha}}, \qquad x \in \mathbb{R}.$$

It can be easily verified that $\mathsf{E}\Lambda_t = t$ for any $t > 0$, but the second moment of Λ_t is infinite due to that the second moment of the random variable V does not exist (and hence, the second moment of the mixed Poisson random sum $S(t)$ with the structural random variable Λ_t does not exist). However, it can be easily seen that

$$\frac{\Lambda_t - t}{\sqrt{t}} = \max\{-\sqrt{t},\, V\} + \frac{1}{2t^{(\alpha+1)/2}}\Big(\frac{2\alpha+1}{\alpha}\sqrt{t} - 1\Big) \Longrightarrow V$$

as $t \to \infty$. This case is an illustrative example of an interesting and non-trivial fact: unlike the classical summation theory, for sums with a random number of summands (in particular, for mixed Poisson random sums) with infinite variances the existence of non-trivial weak limits is possible under the normalization of order $t^{1/2}$ which is «standard» in the classical theory only for sums with finite variances.

The authors have the pleasure to express their gratitude to Margarita Gaponova who carried out the supplementary computations resulting in lemma 6.

R E F E R E N C E S

1. S. V. Artyukhov and V. Yu. Korolev. Estimates of the rate of convergence of the distributions of compound doubly stochastic Poisson processes with non-zero mean to shift mixtures of normal laws. – *Surveys in Industrial and Applied Mathematics*, 2008, Vol. 15, No. 6. p. 988–998 (in Russian).

2. V. E. Bening, V. Yu. Korolev and S. Ya. Shorgin. On approximations to generalized Poisson distribution. – *Journal of Mathematical Sciences*, 1997, Vol. 83, No. 3, p. 360–367.

3. V. Bening and V. Korolev. *Generalized Poisson Models and their Applications in Insurance and Finance.* VSP, Utrecht, The Netherlands, 2002.

4. V. Bentkus. *On the asymptotical behavior of the constant in the Berry–Esseen inequality.* Preprint 91 – 078, Universität Bielefeld, 1991.

5. V. Bentkus. On the asymptotical behavior of the constant in the Berry–Esseen inequality. – *Journal of Theoretical Probability*, 1994, Vol. 2, No, 2, p. 211–224.

6. A. C. Berry. The accuracy of the Gaussian approximation to the sum of independent variables. – *Transactions of the American Mathematical Society*, 1941, Vol. 49, p. 122–136.

7. R. N. Bhattacharya and R. Ranga Rao. *Normal Approximation and Asymptotic Expansions.* Wiley, New York, 1976.

8. P. P. Carr, D. B. Madan and E. C. Chang. The Variance Gamma process and option pricing. – *European Finance Review*, 1998, vol. 2, p. 79–105.

9. P. Delaporte. Un problème de tarification de l'assurance accidents d'automobile examiné par la statistique mathématique. – in: *Trans. 16th Intern. Congress of Actuaries, Brussels*, 1960, Vol. 2, p. 121–135.

10. J. E. A. Dunnage. MR0397839 (53 #1695) 60F05. *Mathematical Reviews on the Web*, American Mathematical Society, 1977.

11. C.-G. Esseen. On the Liapunoff limit of error in the theory of probability. – *Ark. Mat. Astron. Fys.*, 1942, Vol. A28, No. 9, p. 1–19.

12. C.-G. Esseen. A moment inequality with an application to the central limit theorem. – *Skand. Aktuarietidskr.*, 1956, Vol. 39, p. 160–170.

13. M. O. Gaponova and I. G. Shevtsova. Asymptotic estimates of the absolute constant in the Berry–Esseen inequality for distributions with infinite third moments. – *Informatics and Its Applications*, 2009, Vol. 3, No. 4, p. 41-56 (in Russian).

14. S. V. Gavrilenko and V. Yu. Korolev. Convergence rate estimates for mixed Poisson random sums. – In: *Systems and Means of Informatics*, Special Issue. Publishing House of the Institute for Informatics Problems, Russian Academy of Sciences, Moscow, 2006, p. 248–257 (in Russian).

15. B. V. Gnedenko and V. Yu. Korolev. *Random Summation: Limit Theorems and Applications.* CRC Press, Boca Raton, 1996.

16. J. Grandell. *Mixed Poisson Processes.* Chapman and Hall, London, 1997.

17. M. S. Holla. On a Poisson-inverse Gaussian distribution. – *Metrika*, 1967, Vol. 11, p. 115–121.

18. J. O. Irwin. The generalized Waring distribution applied to accident theory. – *Journal of the Royal Statistical Society, Ser. A*, 1968, Vol. 130, p. 205–225.

19. V. Yu. Korolev. A general theorem on the limit behavior of superpositions of independent random processes with applications to Cox processes. – *Journal of Mathematical Sciences*, 1996, Vol. 81, No. 5, p. 2951–2956.

20. V. Yu. Korolev and I. G. Shevtsova. An improvement of the Berry–Esseen inequality. – *Doklady Academii Nauk*, 2010a, to appear (in Russian). English translation: *Doklady Mathematics*, to appear.

21. V. Yu. Korolev and I. G. Shevtsova. A sharpening of the upper estimate of the absolute constant in the Berry–Esseen inequality for mixed Poisson random sums. – *Doklady Academii Nauk*, 2010b, to appear (in Russian). English translation: *Doklady Mathematics*, to appear.

22. V. Yu. Korolev and I. G. Shevtsova. On the upper estimate of the absolute constant in the Berry–Esseen inequality. – *Theory of Probability and its Applications.* 2009, Vol. 54, No. 4, to appear.

23. V. Yu. Korolev and S. Ya. Shorgin. On the absolute constant in the remainder term estimate in the central limit theorem for Poisson random sums. – in: *Probabilistic Methods in Discrete Mathematics*, Proceedings of the Fourth International Petrozavodsk Conference. VSP, Utrecht, 1997, p. 305–308.

24. D. B. Madan and E. Seneta. The variance gamma (V.G.) model for share market return. – *Journal of Business*, 1990, Vol. 63, p. 511–524.

25. R. Michel. On Berry–Esseen results for the compound Poisson distribution. – *Insurance: Mathematics and Economics*, 1993, Vol. 13, No. 1, p. 35–37.

26. V. I. Paulauskas. On a smoothing inequality. – *Lithuanian Mathematical Journal*, 1971, Vol. 11, No. 4, p. 861–866. English translation by American Mathematical Society in *Selected Translations in Mathematical Statistics and Probability*. 1977, Vol. 4, p. 7–12.

27. V. V. Petrov. *Limit Theorems for Sums of Independent Random Variables.* Nauka, Moscow, 1987 (in Russian).

28. V. V. Petrov. *Limit Theorems of Probability Theory. Sequences of Independent Random Variables.* Clarendon Press, Oxford, 1995.

29. H. Prawitz. Limits for a distribution, if the characteristic function is given in a finite domain. – *Scand. AktuarTidskr.*, 1972, p. 138–154.

30. H. Prawitz. *Remainder Term Estimation for Convolution of Identical Components.* – Unpublished manuscript of the lecture given on 16 June, 1972 at the Summer School of the Swedish Statistical Society, 12–21 June, 1972, Löttorp, Sweden.

31. H. Prawitz. Ungleichungen für den absoluten Betrag einer charakteristischen Funktion. – *Skand. AktuarTidskr.*, 1973, No. 1, pp. 11–16.

32. H. Prawitz. Weitere Ungleichungen für den absoluten Betrag einer charakteristischen Funktion. – *Scand. Actuarial J.*, 1975, No. 1, pp. 21–29.

33. H. Prawitz. On the remainder in the central limit theorem. I.– *Scand. Actuarial J.*, 1975, No. 3, p. 145–156.

34. G. V. Rotar. *Some Problems of Reserves Planning.* Candidate Thesis, Central Economical and Mathematical Institute, Moscow, 1972 (in Russian).

35. G. V. Rotar. A problem of storage control. – *Theory of Probability and its Applications.*, 1972, Vol. 17, No. 3, p. 597–599.

36. H. Seal. *Survival Probabilities. The Goal of Risk Theory.* Wiley, Chichester – New York – Brisbane – Toronto, 1978.

37. I. G. Shevtsova. A refinement of the upper estimate of the absolute constant in the Berry–Esseen inequality. – *Theory of Probability and its Applications*, 2006, Vol. 51, No. 3, p. 622–626.

38. I. G. Shevtsova. On the absolute constant in the Berry–Esseen inequality. – In: *The Collection of Papers of Young Scientists of the Faculty of Computational Mathematics and Cybernetics, Moscow State University*, Issue 5. Publishing House of the Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, 2008, p. 101–110 (in Russian).

39. I. G. Shevtsova. The lower asymptotically exact constant in the central limit theorem. – *Doklady Academii Nauk*, 2010a, Vol. 430, No. 4, to appear (in Russian). English translation: *Doklady Mathematics*, to appear.

40. I. G. Shevtsova. On a smoothing inequality. – *Doklady Academii Nauk*, 2010b, Vol. 430, No. 5, to appear (in Russian). English translation: *Doklady Mathematics*, to appear.

41. I. S. Shiganov On a refinement of the upper constant in the remainder term of the central limit theorem. – In: *Stability Problems for Stochastic Models. Proceedings of the Seminar.* Publishing House of the Institute for Systems Studies, Moscow, 1982, p. 109–115 (in Russian).

42. I. S. Shiganov. Refinement of the upper bound of the constant in the central limit theorem. – *Journal of Soviet Mathematics*, 1986, Vol. 35, p. 2545–2550.

43. H. S. Sichel. On a family of discrete distributions particular suited to represent long tailed frequency data. – in: *Proceedings of the 3rd Symposium on Mathematical Statistics*. Ed. by N. F. Laubscher. CSIR, Pretoria, 1971, p. 51–97.

44. I. S. Tyurin. On the convergence rate in Lyapunov's theorem. – *Theory of Probability and its Applications*, to appear.

45. I. S. Tyurin. On the accuracy of the Gaussian approximation. – *Doklady Academii Nauk*, 2009, Vol. 429, No. 3, p. 312–316 (in Russian). English translation: *Doklady Mathematics*, 2009, Vol. 80, No. 3, to appear.

46. I. S. Tyurin. Refinement of the upper bounds of the constants in Lyapunov's theorem. – *Russian Mathematical Surveys*, to appear.

47. I. Tyurin. New estimates of the convergence rate in the Lyapunov theorem. – arXiv:0912.0726v1, 3 December, 2009.

48. J. D. Vaaler. Some extremal functions in Fourier analysis. – *Bulletin of the American Mathematical Society (New Series)*, 1985, Vol. 12, No. 2, p. 183–216.

49. P. van Beek. An application of Fourier methods to the problem of sharpening the Berry–Esseen inequality. – *Z. Wahrsch. verw. Geb.*, 1972, Bd. 23, S. 187–196.

50. R. von Chossy and G. Rappl. Some approximation methods for the distribution of random sums. – *Insurance: Mathematics and Economics*. 1983, Vol. 2, p. 251–270.

51. G. E. Willmot. The Poisson–inverse Gaussian distribution as an alternative to the negative binomial. – *Scandinavian Actuarial Journal*, 1987, p. 113-127.

52. V. M. Zolotarev. On closeness of the distributions of two sums of independent random variables. – *Theory of Probability and its Applications*, 1965, Vol. 10, No. 3, p. 519–526.

53. V. M. Zolotarev. An absolute estimate of the remainder term in the central limit theorem, – *Theory of Probability and its Applications*, 1966, Vol. 11, No. 1, p. 108–119.

54. V. M. Zolotarev. A sharpening of the inequality of Berry–Esseen. – *Z. Wahrsch. verw. Geb.*, 1967, Bd. 8, S. 332–342.

55. V. M. Zolotarev. Some inequalities in probability theory and their application in sharpening the Lyapunov theorem. – *Soviet Math. Dokl.*, 1967, Vol. 8, p. 1427–1430.

56. V. M. Zolotarev. *Modern Theory of Summation of Random Variables*. VSP, Utrecht, The Netherlands, 1997.