Probab. Theory Relat. Fields **111**, 565–583 (1998)

Probability Theory and Related Fields

On signed normal-Poisson approximations

Vydas Čekanavičius

Department of Mathematics, Vilnius University, Naugarduko 24, 2006 Vilnius, Lithuania. e-mail: Vydas.Cekanavicius@maf.vu.lt

Received: 20 November 1996 / Revised version: 5 December 1997

Abstract. For lattice distributions a convolution of two signed Poisson measures proves to be an approximation comparable with the normal law. It enables to get rid of cumbersome summands in asymptotic expansions and to obtain estimates for all Borel sets. Asymptotics can be constructed two-ways: by adding summands to the leading term or by adding summands in its exponent. The choice of approximations is confirmed by the Ibragimov-type necessary and sufficient conditions.

Mathematics Subject Classification (1991): 60 F05

1. Introduction

It would not be an exaggeration to say that the normal approximation plays the leading role in limit theorems. However for continuous distributions it fits much better than for discrete ones. Indeed, the differences in supports rule out a possibility to use total variation or any other stronger than uniform metric. Moreover, even in uniform metric it is impossible to apply the standard Edgeworth expansion, see Gnedenko and Kolmogorov (1968, p. 212). Two approaches are usually used to cope with those problems. The first approach is to replace integral CLT and its refinements by the sum of local theorems, i.e. to replace the normal law by the sum of normal densities – see, for example, Bhattacharya and Ranga Rao (1976), Chap. 5. Unfortu-

---

Key words and phrases: signed Poisson measure; normal approximation; total variation norm

nately such approximation is not infinitely divisible. The second approach is to add cumbersome summands (such as $S(x) = \sum_0^\infty (2\pi v)^{-1} \sin(2\pi v x)$) to the Edgeworth expansion compensating by this the jumps of lattice distributions, see Esseen (1945), Ibragimov and Linnik (1971, p. 100). In fact, both above-mentioned approaches provide much evidence that, for lattice distributions, the normal approximation is not that very natural and some its lattice analogue is needed.

The main purpose of this paper is to show that even in general situations the normal distribution can be replaced by a convolution of Poisson and signed Poisson measure. Such convolution is infinitely divisible and its 'probabilities' can be expressed in Bessel functions. This makes such approximation appropriate for calculations. Moreover, the resemblance in supports enables us to obtain estimates for all Borel sets. Asymptotic expansions can be constructed retaining not only latticeness but also infinite divisibility.

Further we need the following notation. Let be E_a the distribution concentrated at a point a, $E \equiv E_0$. Products and powers of measures are defined in the convolution sense: $FG = F * G$, $F^n = F^{*n}$, $F^0 = E$. For any signed measure of bounded variation W we denote by $W(x) = W\{(-\infty, x)\}$ – the analogue of the distribution function, by $\exp\{W\} = \sum_{k=0}^\infty W^k/k!$ – its exponential measure, by $|W| = \sup_x |W\{(-\infty, x)\}| = \sup_x |W(x)|$ - the analogue of the uniform distance, by $\|W\|$ – the total variation norm of W and by $\widehat{W}(t) = \int_{-\infty}^\infty \exp\{\mathrm{i}t\,x\}\, W\{dx\}$ – its Fourier–Stieltjes transform. Note that $\widehat{\exp}\{W\}(t) = \exp\{\widehat{W}(t)\}$ and $\widehat{E}(t) = 1$.

Let W be concentrated on integers. We denote l_r, $1 \leqslant r < \infty$ (we use the notation l_r instead of a more common l_p) metrics by

$$|W|_r = \left(\sum_{k=-\infty}^{\infty} |W\{k\}|^r \right)^{1/r} . \tag{1.1}$$

Evidently for lattice W we have $\|W\| = |W|_1$. However, understanding the utmost importance of the total variation norm we also use the first notation. It is a well-known fact that the total variation is equivalent to the total variation distance, i.e.,

$$\|W\|/2 \leqslant \sup_B |W\{B\}| \leqslant \|W\| , \tag{1.2}$$

where the supremum is taken over all Borel sets. Note that

$$|W| \leqslant \|W\|, \quad |WV| \leqslant |W|\,\|V\| \tag{1.3}$$

for any two finite measures.

The local distance we denote by

$$|W|_\infty = \sup_k |W\{k\}| \ . \qquad (1.4)$$

Definition. *Let $\lambda \in \mathbb{R}$ and let F be a distribution. Then* $\exp\{\lambda(F - E)\}$ *is called signed compound Poisson* (*SCP*) *measure. A special case of SCP measure, when $F = E_a$* (*i.e. when we have a measure* $\exp\{\lambda(E_a - E)\}$) *is called the signed Poisson* (*SP*) *measure.*

It must be noted that, for $\lambda < 0$, SCP is not a distribution but rather a signed measure. It remains, however, infinitely divisible and always has a finite variation.

In probability theory the elements of signed approximations are not rare. Thus, for example, Le Cam (1960) used a signed compound Binomial approximation. Properties of SCP measures are discussed in Cuppens (1975). However, the turning point in SCP theory were the papers of Presman (1983) and Kornya (1983). Since then SCP measures were applied in actuarial mathematics by Hipp (1986), Dhaene and De Pril (1994), in statistics by Kruopis (1986a), in limit theorems by Kruopis (1986b), Čekanavičius (1991, 1996, 1997), and in probabilistic number theory by Šiaulys and Čekanavičius (1989). See also references in the aforementioned papers.

In this paper we consider one of the most classical situations: a sequence of lattice identically distributed random variables. For our purposes it is more convenient to use distributional notation, thus we usualy write convolutions instead of sums.

Let us assume that ξ has a distribution F and

F does not depend on n, (1.5)

F is concentrated on integers with the greatest common divisor equal to 1, (L_1)

ξ has finite variance and $E\xi = \mu$, $D\xi = \sigma^2$. (1.6)

Note that (L_1) means that ξ is nondegenerate. In fact, any lattice variable can be normed by its maximum span and suitably centered to fulfill (L_1). Thus, all results of this paper can be reformulated for any lattice distribution in the obvious manner.

Further on we use the same notation $C(\cdot)$ for all positive constants depending only on the indicated argument. For example, $C(F)$ is used for all constants depending on F. By θ we denote all quantities satisfying $|\theta| \leqslant C(F)$.

Now we shall introduce the main approximation of this paper. Set

$$D = \exp\left\{\frac{(\sigma^2+\mu)}{2}(E_1 - E) + \frac{(\sigma^2-\mu)}{2}(E_{-1} - E)\right\} \ . \qquad (1.7)$$

Obviously,

$$\widehat{D}(t) = \exp\left\{\frac{(\sigma^2+\mu)}{2}(\mathrm{e}^{\mathrm{i}t}-1)+\frac{(\sigma^2-\mu)}{2}(\mathrm{e}^{-\mathrm{i}t}-1)\right\} . \tag{1.8}$$

If $|\mu| \leqslant \sigma^2$ then D is the convolution of two Poisson distributions, but if $|\mu| > \sigma^2$ then D is the convolution of Poisson law and signed Poisson measure, i.e. it becomes a signed measure. This, however is not a serious fault because any asymptotic expansion is not a distribution. In this sense D is nothing but a Poisson distribution with one member of asymptotics in the exponent.

Approximation D was introduced by Kruopis (1986a,b) who used factorial moments rather than mean and variance. He also proposed to call it the normal – Poisson approximation. Indeed, D is Poisson structured and, for all $t \in \mathbb{R}$,

$$\widehat{D}(t) = \exp\left\{\mathrm{i}t\mu - \frac{\sigma^2 t^2}{2} + \theta|t|^3\right\}, \qquad |\theta| \leqslant C(F) . \tag{1.9}$$

Though Kruopis (1986b) remarked that D^n is uniformly close to F^n even in the general case, he thoroughly explored the Bernoulli case only. His other results employ only the uniform distance and the remainder terms have unestimated summands. In comparison to Kruopis (1986b) we use much weaker assumptions and stronger metrics. Our approach to the asymptotics of D also differs from that used by Kruopis.

The structure of this paper is the following. In Section 2 we introduce asymptotic expansions. In Section 3 we show that those expansions are close to F^n in l_r metrics, provided F has a sufficient number of moments. In Section 4 we explore Ibragimov's necessary and sufficient conditions and prove that they are as much natural for approximation of lattice distributions by D^n, as they are natural for approximation of continuous distributions by the normal law. In concluding remarks we discuss some possible extensions and refinements of the obtained results.

2. Construction of asymptotics for *D*

In this Section we discuss three possible ways of construction of asymptotics. The first and the most general one is so-called Bergström expansion, see Bergström (1951). For its construction no assumptions are needed. Set

$$B_s = D^n + \sum_{j=1}^{s-1}\binom{n}{j}D^{n-j}(F-D)^j . \tag{2.1}$$

From the practical point of view it is more convenient to use infinitely divisible measures whenever we encounter n-fold convolutions. Note that B_s has n-order convolutions only of the infinitely divisible measure D. Meanwhile the largest power of F in (2.1) equals $s-1$. The second advantage is that B_s is a *long* expansion independently of the number of finite moments of F. Bergström expansion does not depend on dimension and can be used in general spaces, see Bentkus (1984), Nagaev and Chebotarev (1989). However, if some number of finite moments exists, it is more reasonable to use a simpler structured expansion.

Let Γ_k be k'th cumulant of F, $(k=3,4,\ldots)$ and let A_k be k'th moment of F, i.e.,

$$\widehat{F}(t) = 1 + \mathrm{i}t\mu + \sum_{k=2}^{s-1} A_k \frac{(\mathrm{i}t)^k}{k!} + \mathrm{o}(|\,t\,|^{s-1}), \quad \text{as } t \to 0 \ , \tag{2.2}$$

$$\ln \widehat{F}(t) = \mathrm{i}t\mu - \frac{\sigma^2 t^2}{2} + \sum_{k=3}^{s-1} \Gamma_k \frac{(\mathrm{i}t)^k}{k!} + \mathrm{o}(|\,t\,|^{s-1}), \quad \text{as } t \to 0 \ . \tag{2.3}$$

In Čekanavičius (1997) we proposed to construct asymptotics in the exponent for the normal distribution and also for a very special case of SP. Now we shall use the same ideas for the construction of asymptotics for D. Note that, for all $t \in \mathbb{R}$,

$$|\widehat{D}(t)| = \exp\{-\sigma^2(1-\cos t)\} = \exp\{-2\sigma^2 \sin^2(t/2)\} \ . \tag{2.4}$$

The main idea is to replace in (2.3) all the powers of t^k by the powers of $(\mathrm{e}^{\mathrm{i}ta}-1)^k$ retaining some analogue of (2.4). The main advantage of $(\mathrm{e}^{\mathrm{i}ta}-1)$ with respect to t is that, for all $t \in \mathbb{R}$, and integer a it satisfies two inequalities simultaneously:

$$|\,\mathrm{e}^{\mathrm{i}ta}-1\,| \leqslant 2, \quad \text{and} \quad |\,\mathrm{e}^{\mathrm{i}ta}-1\,|^2 \leqslant a^2 4\sin^2(t/2) \ . \tag{2.5}$$

Asymptotics in the exponent are constructed step by step. The algorithm is the following. In the first step we replace $\mathrm{i}t\mu - \sigma^2t^2/2$ by $(\sigma^2+\mu)(\mathrm{e}^{\mathrm{i}t}-1)/2 + (\sigma^2-\mu)(\mathrm{e}^{-\mathrm{i}t}-1)/2$ and obtain

$$\begin{aligned}\mathrm{i}t\mu - \frac{\sigma^2t^2}{2} + \sum_{k=3}^{s-1}\Gamma_k\frac{(\mathrm{i}t)^k}{k!} &= \frac{(\sigma^2+\mu)}{2}(\mathrm{e}^{\mathrm{i}t}-1) + \frac{(\sigma^2-\mu)}{2}(\mathrm{e}^{-\mathrm{i}t}-1)\\ &\quad + \sum_{k=3}^{s-1}\left(\Gamma_k - \frac{(\sigma^2+\mu)}{2} - (-1)^k\frac{(\sigma^2-\mu)}{2}\right)\frac{(\mathrm{i}t)^k}{k!}\\ &\quad + \theta|\,t\,|^s \ . \end{aligned}\tag{2.6}$$

Set $\omega_3 = \Gamma_3 - \mu$. Now we shall replace $\omega_3(\mathrm{i}t)^3/6$ by $\lambda_3(\exp\{\mathrm{i}ta(3)\} - 1)^3$, where λ_3 has the same sign as ω_3 and satisfies the equation

$$a^3(3)\lambda_3 = \omega_3/6, \quad \text{where } a(3) = \left[\frac{|\,\omega_3\,|}{\sigma^2}2^5\right] + 1\ .$$

Here $[\cdot]$ means the integer part.

Thus, we have that, for some coefficients $\omega_j, j = 4, 5, \ldots$ and for all $t \in \mathbb{R}$,

$$\lambda_3(\mathrm{e}^{\mathrm{i}ta(3)} - 1)^3 = \omega_3\frac{(\mathrm{i}t)^3}{6} + \sum_{k=4}^{s-1}\omega_k\frac{(\mathrm{i}t)^k}{k!} + \theta|t|^s\ , \tag{2.7}$$

$$\begin{aligned}|\,\lambda_3(\mathrm{e}^{\mathrm{i}ta(3)} - 1)^3\,| &\leqslant 2|\,\lambda_3\,||\,\mathrm{e}^{\mathrm{i}ta(3)} - 1\,|^2 \leqslant 8\sin^2(t/2)|\,\lambda_3\,|a^3(3)/a(3)\\ &\leqslant 2^{-2}\sigma^2\sin^2(t/2)\ .\end{aligned} \tag{2.8}$$

The next step is to replace $(\Gamma_4 - \sigma^2 - \omega_4)(\mathrm{i}t)^4/4!$ by some $\lambda_4(\exp\{\mathrm{i}ta(4)\} - 1)^4$, where $a(4)$ is a natural number and λ_4 satisfies assumptions that lead to the analogue of (2.7) and (2.8). In general the algorithm of replacement of $\omega(\mathrm{i}t)^k/k!$ by $\lambda_k(\exp\{\mathrm{i}ta(k)\} - 1)^k$ is the following. Let λ_k be of the same sign as ω_k and let

$$\lambda_k a^k(k) = \omega/k!\ , \tag{2.9}$$

where

$$a(k) = \left[|\,\omega\,|^{1/(k-2)}\sigma^{-2/(k-2)}2^{(2k-1)/(k-2)}\right] + 1\ . \tag{2.10}$$

After s steps, for all $t \in \mathbb{R}$, we get

$$\begin{aligned}\mathrm{i}t\mu - \frac{\sigma^2t^2}{2} + \sum_{k=3}^{s-1}\Gamma_k\frac{(\mathrm{i}t)^k}{k!} &= \frac{(\sigma^2 + \mu)}{2}(\mathrm{e}^{\mathrm{i}t} - 1) + \frac{(\sigma^2 - \mu)}{2}(\mathrm{e}^{-\mathrm{i}t} - 1)\\ &\quad + \sum_{k=3}^{s-1}\lambda_k(\mathrm{e}^{\mathrm{i}ta(k)} - 1)^k + \theta|\,t\,|^s\ .\end{aligned} \tag{2.11}$$

Moreover, due to the construction

$$\begin{aligned}\sum_{k=3}^{s-1}\left|\lambda_k\left(\mathrm{e}^{\mathrm{i}ta(k)} - 1\right)^k\right| &\leqslant \sum_{k=3}^{s-1}|\lambda_k|2^{k-2}a^2(k)4\sin^2(t/2)\\ &\leqslant \sum_{k=3}^{s-1}2^k\sin^2(t/2)\sigma^2 2^{1-2k} \leqslant \frac{\sigma^2}{2}\sin^2(t/2)\ .\end{aligned} \tag{2.12}$$

Set

$$D_s = \exp\left\{\frac{(\sigma^2+\mu)}{2}(E_1 - E) + \frac{(\sigma^2-\mu)}{2}(E_{-1} - E) + \sum_{k=3}^{s-1} \lambda_k (E_{a(k)} - E)^k\right\} \ . \tag{2.13}$$

Assume that F has a finite absolute moment of order $s \geqslant 3$. By (2.11), (2.12) and the relations between cumulants and moments, for all $t \in \mathbb{R}$, we have

$$|\widehat{D}_s(t)| \leqslant \exp\left\{-\frac{3\sigma^2}{2}\sin^2\frac{t}{2}\right\} \ , \tag{2.14}$$

$$|\widehat{F}(t) - \widehat{D}_s(t)| \leq C(F,s)|t|^s \ . \tag{2.15}$$

After a quite standard calculation we also get that, for all $t \in \mathbb{R}$,

$$|\widehat{F}'(t) - \widehat{D}_s'(t)| \leqslant C(F,s)|t|^{s-1} \tag{2.16}$$

and

$$\left|(\widehat{F}(t)\mathrm{e}^{-\mathrm{i}t\mu})'\right| \leqslant C(F)|t|, \quad |(\widehat{D}_s(t)\mathrm{e}^{-\mathrm{i}t\mu})'| \leqslant C(F,s)|t| \ . \tag{2.17}$$

D_s is both infinitely divisible and lattice. Note that the construction of D_s is quite arbitrary and other λ_k and $a(k)$ can be selected provided the relation (2.11) is fulfilled and the analogue of (2.14) holds. Here we presented just one of the possibilities.

If we want to construct asymptotics by adding summands to D, then we should put down some of the expansion in the exponent. Kruopis (1986b) considered one of such expansions based on factorial moments. Here we shall introduce a different approach based on cumulants. For the sake of brevity we shall use only a short expansion. We can write formally

$$\widehat{F}^n(t) = \exp\{n \ln \widehat{F}(t)\} = \widehat{D}^n(t)\left(1 + n(\Gamma_3 - \mu)\frac{(\mathrm{i}t)^3}{3!} + \cdots\right) \ . \tag{2.18}$$

Now let us replace $(\mathrm{i}t)^3$ by $(\mathrm{e}^{\mathrm{i}t} - 1)^3$. Note that in this case we do not need the analogue of (2.8). Moreover, the replacement is not unique. Indeed, we can use $(1 - \mathrm{e}^{-\mathrm{i}t})(\mathrm{e}^{\mathrm{i}t} - 1)^2$ etc. Set

$$K_3(n) = D^n\left(E + n\frac{\Gamma_3 - \mu}{6}(E_1 - E)^3\right) \ . \tag{2.19}$$

If we want to get a longer expansion then evidently we should take more summands in (2.18) and replace $(\mathrm{i}t)^k$ by some $\lambda_k(\mathrm{e}^{\mathrm{i}t} - 1)^k$ in a manner analogous to the construction of D_s with the only exception that the analogue of (2.14) is unnecessary.

We shall end this Section by formulating important consequences of the conditions (1.5), (L_1) and (1.6). It is easy to check that, if those conditions are satisfied then, for some $0 < \varepsilon \leqslant \pi$,

$$|\widehat{F}(t)| \leqslant \mathrm{e}^{-C(F)t^2}, \qquad \text{if } |t| \leqslant \varepsilon \ . \tag{2.20}$$

Moreover, for any $0 < \varepsilon \leqslant \pi$ and all $\varepsilon \leqslant |t| \leqslant \pi$,

$$|\widehat{F}(t)|^n \leqslant \mathrm{e}^{-nC(F,\varepsilon)}, \tag{2.21}$$

see Gnedenko and Kolmogorov (1968, p. 234). Note also that from (2.14) we get that, for $|t| \leqslant \pi$,

$$|\widehat{D}_s(t)| \leqslant \mathrm{e}^{-C(F,s)t^2} \ . \tag{2.22}$$

3. Estimates in ℓ_r metric

Throughout this Section we assume that $1 \leqslant r < \infty$ is a fixed number, s is a natural number and by writing $\mathrm{O}(n^{-s})$ we consider $n \to \infty$.

As we have mentioned previously, Bergström expansions require a very little amount of information. The following long expansion requires only the existence of three absolute moments.

Theorem 3.1. *Let F satisfy conditions* (1.5), (1.6) *and* (L_1) *and let*

$$\int_{-\infty}^{\infty} |x|^3 F\{dx\} = \beta < \infty \ .$$

Then, for any fixed $s \geqslant 1$,

$$|F^n - B_s|_r = \mathrm{O}\left(n^{-s/2-(r-1)/(2r)}\right) \ , \tag{3.1}$$

$$|F^n - B_s|_\infty = \mathrm{O}\left(n^{-(s+1)/2}\right) \ . \tag{3.2}$$

Corollary 3.1. *Under the conditions of Theorem* 3.1

$$\|F^n - B_s\| = \mathrm{O}\left(n^{-s/2}\right) \ . \tag{3.3}$$

If F has the s'th finite absolute moment we can apply the approximation D_s.

Theorem 3.2. *Let F satisfy* (1.5), (1.6) *and* (L_1) *and have finite s'th absolute moments. Then*

$$|F^n - D_s^n|_r = \mathrm{O}\left(n^{-(s-2)/2-(r-1)/(2r)}\right) \ , \tag{3.4}$$

$$|F^n - D_s^n|_\infty = \mathrm{O}\Big(n^{-(s-1)/2}\Big) \ . \tag{3.5}$$

Corollary 3.2. *Under the conditions of Theorem* 3.2

$$\|F^n - D_s^n\| = \mathrm{O}\Big(n^{-(s-2)/2}\Big) \ . \tag{3.6}$$

We can apply $K_3(n)$ too.

Theorem 3.3. *Let F satisfy* (1.5), (1.6) and (L_1) *and have finite third absolute moment. Then*

$$|F^n - K_3(n)|_r = \mathrm{O}\big(n^{-1-(r-1)/(2r)}\big) \ , \tag{3.7}$$

$$|F^n - K_3(n)\,|_\infty \ = \ \mathrm{O}\big(n^{-3/2}\big) \ , \tag{3.8}$$

$$\|F^n - K_3(n)\| = \mathrm{O}\big(n^{-1}\big) \ . \tag{3.9}$$

Before proving theorems we shall discuss some aspects of their application. In all cases the fast Fourier transform can be used for calculation of ‘probabilities’. However it is evident that, in fact, B_s is not as convenient as the remaining two approximations, because in B_s the initial distribution F is used (not only its moments). Thus B_s to some extent can be viewed as an intermediate approximation. For $K_3(n)$ it is possible to use some recursion algorithm. Indeed, let

$$P(m) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}tm}\widehat{D}^n(t)\,dt \ .$$

Then $P(m)$ satisfies a simple recursion formula

$$(\mu - \sigma^2)P(m+1) = 2mP(m) + (\sigma^2 + \mu)P(m-1) \ . \tag{3.10}$$

But $K_3(n)$ can be expressed in $P(m)$’s as follows

$$\begin{aligned} K_3(n)\{m\} = \ & P(m) + n\frac{(\Gamma_3 - \mu)}{6} \\ & \times (P(m-3) - 3P(m-2) + 3P(m-1) - P(m)) \ . \end{aligned} \tag{3.11}$$

Now for calculation (3.10) can be applied. Note that $P(0), P(1), P(-1)$ can be calculated with a given accuracy directly from the inversion formula. Besides $P(m)$ can be expressed in Bessel functions. For example, let $\sigma^2 + \mu \geqslant 0$ and $\sigma^2 - \mu < 0$. Then

$$P(m) = \mathrm{e}^{-\sigma^2}(\sigma^2 + \mu)^{m/2}|\sigma^2 - \mu|^{-m/2}J_m\Big(\sqrt{|\sigma^4 - \mu^2|}\Big) \ ,$$

where J_m is the Bessel function of the first kind. For other values of μ and σ analogous formulas can be obtained.

Now we pass onto the proofs of theorems. The following auxiliary result is needed.

Lemma 3.1. *Let Δ be a measure of finite variation concentrated on integers. Then*

$$|\Delta|_r \leqslant \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |\widehat{\Delta}(t)|^{r/(r-1)}\, dt \right)^{(r-1)/r}, \quad \text{for } r \geqslant 2 \tag{3.12}$$

$$|\Delta|_\infty \leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} |\widehat{\Delta}(t)|\, dt\ , \tag{3.13}$$

$$|\Delta| \leqslant \frac{1}{4} \int_{-\pi}^{\pi} \frac{|\widehat{\Delta}(t)|}{|t|}\, dt\ . \tag{3.14}$$

If, in addition, $\sum_k |k| |\Delta\{k\}| < \infty$ *then for all* $a \in \mathbb{R}$, $b > 0$ and $1 \leqslant r < 2$

$$|\Delta|_r \leqslant \left(1 + b\pi\sqrt{r/(2-r)}\right)^{(2-r)/(2r)} \times \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |\widehat{\Delta}(t)|^2 + \frac{2-r}{rb^2} \left| \left(\mathrm{e}^{-\mathrm{i}ta}\widehat{\Delta}(t)\right)' \right|^2 dt \right)^{1/2}. \tag{3.15}$$

The relation (3.13) follows directly from the inversion formula. The relation (3.14) is a well-known Tsaregradskii's inequality, see Tsaregradskii (1958). Proof of other relations can be found in Šiaulys and Čekanavičius (1988).

Proof of Theorem 3.1. For the sake of brevity we write $\widehat{F}, \widehat{D}, B_s, C$ instead of $\widehat{F}(t), \widehat{D}(t), \widehat{B}_s(t)$, $C(F, s, r, \varepsilon)$. Note that C is used for different constants. By the Bergström (1951) identity we have

$$|\widehat{F}^n - \widehat{B}_s| = \left| \sum_{j=s}^{n} \binom{j-1}{s-1} \widehat{F}^{n-j} (\widehat{F} - \widehat{D})^s \widehat{D}^{j-s} \right| . \tag{3.16}$$

Let $|t| \leqslant \varepsilon$, where ε is defined by (2.20). Then

$$|\widehat{F}^n - \widehat{B}_s| \leqslant \sum_{j=s}^{n} \binom{j-1}{s-1} \mathrm{e}^{-(n-s)Ct^2} |\widehat{F} - \widehat{D}|^s$$

$$\leqslant Cn^s \mathrm{e}^{-nCt^2} |\, t\,|^3 \leqslant Cn^{-s/2} \mathrm{e}^{-nCt^2} \ . \tag{3.17}$$

By (2.21) and (2.22) we obtain, for $\varepsilon \leqslant |\, t\,| \leqslant \pi$,

$$|\,\widehat{F} - \widehat{B}_s\,| \leqslant Cn^s \mathrm{e}^{-nC} = \mathrm{O}\big(n^{-(s+1)/2}\big), \qquad \text{as } n \to \infty \ . \tag{3.18}$$

From (3.17), (3.18) and (3.13) we get (3.2). Quite analogously we prove (3.1) for $r \geqslant 2$. Now let $1 \leqslant r < 2$. Then, taking $a = (n - s)\mu$, for all $|\, t\,| \leqslant \varepsilon$ by (2.17) and (2.18) we obtain

$$\begin{aligned} |\,\big((\widehat{F}^n - \widehat{B}_s)\mathrm{e}^{-\mathrm{i}ta}\big)'\,| &\leqslant \sum_{j=s}^{n} \binom{j-1}{s-1} \big|\big((\widehat{F}\mathrm{e}^{-\mathrm{i}t\mu})^{n-j}(\widehat{D}\mathrm{e}^{-\mathrm{i}t\mu})^{j-s}(\widehat{F} - \widehat{D})^s\big)'\big| \\ &\leqslant C\mathrm{e}^{-nCt^2} n^s \big(n|\, t\,|^{1+3s} + |\, t\,|^{3(s-1)+2}\big) \\ &\leqslant C\mathrm{e}^{-nCt^2} n^{-(s-1)/2} \ . \end{aligned} \tag{3.19}$$

From (2.21) and (2.22) we get that, for $\varepsilon \leqslant |\, t\,| \leqslant \pi$,

$$\big|\big(\mathrm{e}^{-\mathrm{i}ta}(\widehat{F}^n - \widehat{B}_s)\big)'\big|^2 = \mathrm{O}\big(n^{-s/2-(r-1)/2r}\big) \ . \tag{3.20}$$

Putting $b = \sqrt{n}$ in (3.15) and taking into account (3.17)–(3.20) we get the assertion of the theorem. □

Proof of Theorem 3.2. From (2.21) and (2.22) we see that it suffices to get the estimates only for $|\, t\,| \leqslant \varepsilon$. But, for $|\, t\,| \leqslant \varepsilon$,

$$\begin{aligned} |\,\widehat{F}^n - \widehat{D}_s^n\,| &\leqslant n \max(|\,\widehat{F}\,|^{n-1}, |\,\widehat{D}_s\,|^{n-1}) |\,\widehat{F} - \widehat{D}_s\,| \leqslant Cn\mathrm{e}^{-nCt^2} |\, t\,|^s \\ &\leqslant C\mathrm{e}^{-nCt^2} n^{-(s-2)/2} \end{aligned} \tag{3.21}$$

and

$$\begin{aligned} |\,(\mathrm{e}^{-\mathrm{i}t(n-1)\mu}(\widehat{F}^n - \widehat{D}_s^n))'\,| &\leqslant \sum_{j=1}^{n} |\,\big((\widehat{F}\mathrm{e}^{-\mathrm{i}t\mu})^{n-j}(\widehat{D}_s\mathrm{e}^{-\mathrm{i}t\mu})^{j-1}(\widehat{F} - \widehat{D}_s)\big)'\,| \\ &\leqslant Cn(n|\, t\,|^{1+s} + |\, t\,|^{s-1})\mathrm{e}^{-nCt^2} \\ &\leqslant C\mathrm{e}^{-nCt^2} n^{-(s-3)/2} \ . \end{aligned}$$

Applying Lemma 3.1 with $a = (n - 1)\mu$, $b = \sqrt{n}$ we get the assertion of the theorem. □

Proof of Theorem 3.3. By the properties of metrics we have

$$\begin{aligned} |\, F^n - K_3(n)\,|_r &\leqslant |\, F^n - B_2\,|_r + |\, B_2 - K_3(n)\,|_r \ , \\ |\, B_2 - K_3(n)\,|_r &\leqslant n|\, D^{n-1}(F - D)(E - D)\,|_r \\ &+ n|\, D^n(F - D - \frac{\Gamma_3 - \mu}{6}(E_1 - E)^3)\,|_r \ . \end{aligned}$$

Now we should apply Theorem 3.1 and Lemma 3.1 (once with $b = \sqrt{n}$ and $a = (n-1)\mu$ and once with $b = \sqrt{n}$ and $a = n\mu$). As all calculations are quite analogous to the previous ones we omit the remaining part of the proof. $\square$

Remark. Proof of the Theorem 3.3 demonstrates one of the properties of Bergström expansions, i.e., their usefulness as the intermediate results.

4. Necessary and sufficient conditions

Ibragimov was the first who established necessary and sufficient conditions for the rate of convergence to the normal distribution, see Ibragimov (1966, 1967). Later Ibragimov's conditions were reformulated and generalized in many directions, see, Bikelis (1972), Rozovskii (1978), Michel (1983) and references therein. However in this paper we shall stick to the original Ibragimov version, because it clearly demonstrates that the most appropriate case for the normal approximation is the case when F satisfies Cramer's (C) condition, i.e.,

$$\limsup_{|t|\to\infty} |\widehat{F}(t)| < 1 \ . \tag{C}$$

For lattice distributions the lattice structured approximation is more natural. To this goal we shall show that D^n_{s+3} is close to F^n for all Borel sets. Let ξ have a distribution F which satisfies (1.5), (1.6) and has finite cumulants $\Gamma_3, \Gamma_4, \dots, \Gamma_{s+2}$. Let F_η correspond to $\eta = (\xi - \mu)/\sigma$. Cumulants of F_η we denote by $\mu_1, \mu_2, \dots, \mu_{s+2}$. Let D_{s+3} be constructed as in (2.13), i.e. D_{s+3} has the same form as in (2.13) and $\lambda_j, a(j)$ are such that

$$\widehat{D}_{s+3}(t/\sigma) = \mathrm{e}^{\mathrm{i}t\mu/\sigma} h_s(t) \exp\{\theta |t|^{s+3}\} \ , \tag{4.1}$$

where

$$h_s(t) = \exp\left\{ -\frac{t^2}{2} + \sum_{k=3}^{s+2} \frac{(\mathrm{i}t)^k}{k!}\mu_k \right\} \ . \tag{4.2}$$

Besides, for all $t \in \mathbb{R}$

$$|\widehat{D}_{s+3}(t)| \leqslant \exp\left\{-\frac{3\sigma^2}{2}\sin^2\frac{t}{2}\right\} \ . \tag{4.3}$$

Now we can formulate the main result of this Section.

Theorem 4.1. *For the relation*

$$\| F^n - D^n_{s+3} \| = \mathrm{O}(n^{-(s+\delta)/2}) \ , \tag{4.4}$$

$s = 1, 2, \ldots, 0 < \delta < 1$ *to hold it is necessary* (*and, for F satisfying* (L_1), *sufficient*) *that the condition*

$$\int\limits_{|x-\mu|>z} |x-\mu|^{s+2} dF = \mathrm{O}(z^{-\delta}), \qquad z \to \infty \ , \tag{4.5}$$

be satisfied.

For the relation (4.4) *to hold with* $\delta = 1$ *it is necessary* (*and, for F satisfying* (L_1) *sufficient*) *that the conditions* (4.5) *and*

$$\int\limits_{\mu-z}^{\mu+z} (x-\mu)^{s+3}\, dF = \mathrm{O}(1), \ \ z \to \infty \tag{4.6}$$

be satisfied.

Conditions (4.5)–(4.6) are the well-known Ibragimov conditions for the uniform normal approximation, see Ibragimov (1966). In fact, we shall prove a stronger proposition, namely we shall show that (4.5)–(4.6) are necessary for

$$|F^n - D^n_{s+3}| = \mathrm{O}(n^{-(s+\delta)/2}) \ , \tag{4.7}$$

and sufficient for (4.4). By the properties of distances

$$|F^n - D^n_{s+3}| \leqslant \| F^n - D^n_{s+3} \| \ . \tag{4.8}$$

Therefore, (4.7) will follow immediately from (4.4) and (4.8). Moreover, from (4.8) we shall get that (4.5) and (4.6) are necessary and sufficient for (4.7).

Proof of Theorem 4.1. Necessary part. Ibragimov (1966, 1967) proved that (4.5)–(4.6) are necessary and sufficient for the closeness of the normed sum to the standard normal distribution. We shall prove the necessity of (4.5) and (4.6) by induction. Let $s = 0$. Then $D_{s+3} = D_3 = D$. From Ibragimov's (1966) paper we have that (4.5) and (4.6) are necessary for

$$|F^n - \Phi(n\mu, \sigma\sqrt{n})| = \mathrm{O}(n^{-\delta/2}) \ , \tag{4.9}$$

where $\Phi(n\mu, \sigma\sqrt{n})$ is the normal distribution with the mean $n\mu$ and variance $n\sigma^2$. Applying variant of Esseen's smoothing lemma from Petrov (1975, Chap. 5, Th.1) we get

$$|D^n - \Phi(n\mu, \sigma\sqrt{n})| \leqslant C \int\limits_{-\pi}^{\pi} |\widehat{D}(t) - \exp\{\mathrm{i}tn\mu - n\sigma^2 t^2/2\}| \frac{dt}{|t|} + C$$

$$\times \int_{|y|\leqslant C} |D^n(x+y) - D^n(x)|\, dy \ . \tag{4.10}$$

By (3.13) and (2.22)

$$|D^n(x+y) - D^n(x)| \leqslant C(|y|+1) \max_k |D^n\{k\}|$$
$$\leqslant \frac{C(|y|+1)}{2\pi} \int_{-\pi}^{\pi} |\widehat{D}^n(t)|\, dt \leqslant C(|y|+1)n^{-1/2} \ . \tag{4.11}$$

On the other hand, the first integral in (4.10) is also majorized by $C(F)n^{-1/2}$. Hence

$$|D^n - \Phi(n\mu, \sigma\sqrt{n})| = \mathrm{O}(n^{-1/2}) \ . \tag{4.12}$$

Taking into account (4.12) and (4.8) we see that from (4.4) it follows (4.7) and, subsequently, (4.9). The conditions (4.5)–(4.6) are necessary for (4.9). The necessary part, for $s = 0$, is proved. Now let (4.4)–(4.6) be satisfied for some s, i.e., let them be necessary for

$$|F^n - D^n_{s+3}| = \mathrm{O}(n^{-(s+\delta)/2}) \ . \tag{4.13}$$

We shall prove that then (4.5), (4.6)(with $s+1$) are necessary for

$$|F^n - D^n_{s+4}| = \mathrm{O}(n^{-(s+1+\delta)/2}) \ . \tag{4.14}$$

Just as in Ibragimov (1967) set

$$A(t) = \begin{cases} t(1-t)h^n(t/\sqrt{n}), & t \in [0,1] \\ 0, & t \notin [0,1] \end{cases} \ .$$

Let $\tilde{A}(x)$ be the Fourier transform of $A(t)$. Then by the properties of uniform distance

$$|F^n - D^n_{s+4}| = \sup_x |F^n_\eta(\sqrt{n}x) - E_{-\mu n/\sigma\sqrt{n}} D^n_{s+4}(\sqrt{n}\sigma x)| = \sup_x |\Delta(x)|,$$
say .

By Parseval's identity

$$\left| \int_0^1 \frac{\widehat{\Delta}(t)}{\mathrm{i}t} A(t)\, dt \right| \leqslant \int_{-\infty}^{\infty} |\Delta(x)||\tilde{A}(x)|\, dx = \mathrm{O}(n^{-(s+1+\delta)/2}) \ . \tag{4.15}$$

But by (4.1)

$$\widehat{\Delta}(t) = \widehat{F}^n_\eta(t/\sqrt{n}) - h_{s+1}(t/\sqrt{n}) \exp\{\theta|t|^{s+4} n^{-(s+2)/2}\} \ . \tag{4.16}$$

Ibragimov (1967) proved that, for $|t| \leqslant 1$,

$$g_{s+1}(t) = h^n_{s+1}(t/\sqrt{n}) \exp\left\{ \mathrm{O}\left(|t|^{s+2} n^{-(s+2)/2} \right) \right\}, \qquad n \to \infty \ . \tag{4.17}$$

From (4.15)–(4.17) it is not difficult to obtain

$$\left| \int_0^1 (\widehat{F}^n_\eta(t/\sqrt{n}) - g_{s+1}(t)) \frac{A(t)}{\mathrm{i}t} \, dt \right| = \mathrm{O}(n^{-(s+1+\delta)/2}) \ . \tag{4.18}$$

But Ibragimov (1967) proved that (4.18) and necessity of (4.5)–(4.6) for s imply the necessity of (4.5)–(4.6) for $s+1$. Thus, the necessary part is proved.

Sufficient part. Ibragimov (1967) proved that (4.5)–(4.6) imply

$$\widehat{F}_\eta(t) = h_s(t) \exp\left\{ \mathrm{O}\left(|t|^{s+2+\delta} \right) \right\} \qquad t \to 0, \quad 0 < \delta \leqslant 1 \ . \tag{4.19}$$

Therefore, for some ε_1 by (4.1) and (4.19) we get

$$\widehat{F}(t) = \mathrm{e}^{\mathrm{i}t\mu} h_s(t\sigma) \exp\{\theta |t|^{s+2+\delta}\}, \qquad \text{for } |t| \leqslant \varepsilon_1 \tag{4.20}$$

$$\widehat{D}_{s+3}(t) = \mathrm{e}^{\mathrm{i}t\mu} h_s(t\sigma) \exp\{\theta |t|^{s+2+\delta}\} \qquad \text{for all } |t| \leqslant \pi \ . \tag{4.21}$$

Let $\beta_1, \beta_2, \ldots$ be the moments of η. Then, just like in Ibragimov (1966), (1967) we get that

$$\begin{aligned} \left| \widehat{F}_\eta(t) - \sum_{k=0}^{s+2} \frac{(\mathrm{i}t)^k}{k!} \beta_k \right| &= \mathrm{O}(|t|^{s+2+\delta}), \\ \left| \widehat{F}'_\eta(t) - \left(\sum_{k=0}^{s+2} \frac{(\mathrm{i}t)^k}{k!} \beta_k \right)' \right| &= \mathrm{O}(|t|^{s+1+\delta}) \ . \end{aligned} \tag{4.22}$$

After quite standard calculation from (4.20)–(4.22) and (4.3) we get that, for some $0 < \varepsilon_2 \leqslant \pi$ and all $|t| \leqslant \varepsilon_2$,

$$\begin{aligned} |\widehat{F}(t) - \widehat{D}_{s+3}(t)| &\leqslant C(F, \varepsilon_3) |t|^{s+2+\delta} \ , \\ |\widehat{F}'(t) - \widehat{D}'_{s+3}(t)| &\leqslant C(F, \varepsilon_3) |t|^{s+1+\delta} \ . \end{aligned}$$

But by (2.21) and (2.22), for $|t| > \varepsilon_2$, all $|\widehat{F}(t)|^n$, $|\widehat{D}_{s+3}(t)|^n$ and their derivatives vanish exponentially. Thus to end the proof it suffices to proceed as in Theorem 3.1. □

Of course, it is also possible to use $K_3(n)$ or other expansions. For example, let us consider $K_3(n)$ and the uniform distance. According to Ibragimov (1967) the conditions (4.5)–(4.6) can be replaced by the condition on characteristic function. We shall use the same approach.

Theorem 4.2. *For the relation*

$$|\,F^n - K_3(n)\,| \;=\; \mathrm{O}\Big(n^{-(1+\delta)/2}\Big), \qquad 0 \;<\; \delta \leqslant 1 \tag{4.23}$$

to hold it is necessary (*and*, *for F satisfying* (L_1), *sufficient*) *that*

$$\widehat{F}(t) = \exp\left\{\mathrm{i}t\mu - \frac{\sigma^2 t^2}{2} + \Gamma_3 \frac{(\mathrm{i}t)^3}{6} + \mathrm{O}\Big(|t|^{3+\delta}\Big)\right\}\ . \tag{4.24}$$

Proof. Ibragimov (1967) proved that from (4.5) and (4.6) it follows (4.24). Now the necessary part can be proved analogously to the necessary part in Theorem 4.1 and the sufficient part follows from (4.24) and (3.14). □

5. Concluding remarks

Asymptotic expansion in the exponent of Section 2 is not the only possible one. In fact, λ_k and $a(k)$ can be chosen in an arbitrary way provided (2.11) and (2.12) hold true. The idea of fitting as much moments as possible lead Rachev and Rüschendorf (1990) to the ‘scaled’ Poisson approximations. However, their approach means that the approximated distribution and the approximating distribution have different maximal spans – and all problems, analogous to those of normal approach, remain.

The estimates of this paper are obtained for the scheme of sequences. If we require in addition that $F\{0\}, F\{1\} > 0$ then it is possible to reformulate all results in the general scheme of series, where the dependence of $C(F)$ on F will be explicit. The main difference in the proofs is the following estimate

$$\begin{aligned}
|\,\widehat{F}(t)\,|^2 &= \left(F\{0\} + \sum_1^\infty F\{j\}\cos(tj)\right)^2 + \left(\sum_1^\infty F\{j\}\sin(tj)\right)^2 \\
&\leqslant F^2\{0\} + 2F\{0\}\sum_1^\infty F\{j\}\cos(tj) + \left(\sum_1^\infty F\{j\}\right)^2 \\
&\leqslant 1 - 2F\{0\}\sum_1^\infty F\{j\}\sin^2(tj/2) \\
&\leqslant \exp\{-2F\{0\}F\{1\}\sin^2(t/2)\},
\end{aligned} \tag{5.1}$$

which holds for all $t \in \mathbb{R}$.

Assuming more of finite moments it is not difficult to get the lower estimates. In this case the Theorem 2 from Šiaulys and Čekanavičius (1988) can be applied.

Example 5.1. Let F, D be as defined in Sections 1–2. Assume that F has a finite fourth moment A_4 and let

$$| \mu + 3\mu\sigma^2 + \mu^3 - A_3 | > 0, \quad \text{where } A_3 = \int x^3 F\{dx\} \ .$$

Directly applying aforementioned theorem we get that for all $n \geqslant 3$

$$| F^n - D^n | \geqslant C(F) n^{-1/2} \ . \tag{5.2}$$

So far in this paper we considered identically distributed summands (i.e. F^n). Now we give one example of approximation of sum with nonidentically distributed summands.

Example 5.2. Let us consider the Wilcoxon signed rank statistic. The well known fact is that this statistic is uniformly close to the normal distribution. Moreover, it can be decomposed in the sum of independent nonidentically distributed two-valued random variables. In the terms of distributions (denoting by F_n the distribution of Wilcoxon statistic) we have

$$F_n = \prod_{j=1}^{n} \Big(\frac{1}{2} E + \frac{1}{2} E_j \Big) = \prod_{j=1}^{n} F(j), \text{ say} \ .$$

Note that

$$\widehat{F}_n(t) = \prod_{j=1}^{n} (1/2 + \mathrm{e}^{\mathrm{i}tj}/2) \ . \tag{5.3}$$

The mean and the variance of F_n respectively are

$$\mu(n) = n(n+1)/4, \qquad \sigma^2(n) = n(n+1)(2n+1)/24 \ . \tag{5.4}$$

Let D_n be defined as in Section 2, i.e.,

$$D_n = \exp\left\{ \frac{\sigma^2(n) + \mu(n)}{2} (E_1 - E) + \frac{\sigma^2(n) - \mu(n)}{2} (E_{-1} - E) \right\} \ . \tag{5.5}$$

We have that $D_n = \prod_1^n D(j)$, where

$$\widehat{D}(j)(t) = \exp\left\{ \frac{(2j + j^2)}{8} (\mathrm{e}^{\mathrm{i}t} - 1) + \frac{(j^2 - 2j)}{8} (\mathrm{e}^{-\mathrm{i}t} - 1) \right\} \ . \tag{5.6}$$

In Quine (1994) it was shown that for $| t | > C_1/n$

$$| \widehat{F}_n(t) | \leqslant \mathrm{e}^{-C_2 n} \ . \tag{5.7}$$

Proceeding just as in Section 3 we establish that for $|t| \leqslant C_1/n$

$$|\widehat{F}(j)(t) - \widehat{D}(j)(t)| \leqslant C(|t| + |t|^3 j^3) \ , \tag{5.8}$$

$$|(\widehat{F}(j)(t)\mathrm{e}^{-\mathrm{i}tj/2})' - (\widehat{D}(j)(t)\mathrm{e}^{-\mathrm{i}tj/2})'| \leqslant C(1 + t^2 j^3) \ , \tag{5.9}$$

$$|(\widehat{F}(j)(t)\mathrm{e}^{-\mathrm{i}tj/2})'|, |(\widehat{D}(j)(t)\mathrm{e}^{-\mathrm{i}tj/2})'| \leqslant C|t|j^2 \ , \tag{5.10}$$

$$\prod_{j=1}^{n} |\widehat{F}(j)(t)| \leqslant C \exp\{-Cn^3t^2\}, \quad \prod_{j=1}^{n} |\widehat{D}(j)(t)| \leqslant C \exp\{-Cn^3t^2\} \ . \tag{5.11}$$

From (5.7)–(5.11) it is not difficult to obtain

$$\| F_n - D_n \| = \mathrm{O}(n^{-1/2}) \ . \tag{5.12}$$

Thus we see that, just as in the identically distributed case, the order of the estimate is the same as for the normal approximation, but (unlike the normal case) it holds for all Borel sets.

Acknowledgement. I thank the referees for useful suggestions which helped to improve the paper.

References

Bentkus, V.: Asymptotic expansions in the CLT in Hilbert space. Lithuan. Math. J. **24**, 210–225 (1984)

Bergström, H.: On asymptotic expansion of probability functions. Skand. Aktuarietidskrift **1**, 1–34 (1951)

Bhattacharya, R.N., Ranga Rao, R.: Normal Approximation and asymptotic expansions. New York: Wiley (1976)

Bikelis, A.: Limit theorems for the sums of independent random variables. (In Russian) Litovsk. Matem. Sb. **4**, 4–14 (1972)

Čekanavičius, V.: Approximation of mixtures of distributions. Lithuan. Math. J. **31**, 243–257 (1991)

Čekanavičius, V.: On multivariate Le Cam theorem and compound Poisson measures. Statist. and Probab. Lett. **28**, 33–39 (1996)

Čekanavičius, V.: Asymptotic expansions in the exponent: a compound Poisson approach. Adv. Appl. Probab. **29**, 374–387 (1997)

Cuppens, R.: Decomposition of multivariate probability. New York, San Francisco, London (1975)

Dhaene, J., De Pril, N.: On a class of approximative computation methods in the individual risk model. Insurance: Math. and Economics. **14**, 181–196 (1994)

Esseen, C.G.: Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. Acta Math. **77**, 1–125 (1945)

Gnedenko, B.V., Kolmogorov, A.N.: Limit distributions for sums of independent random variables. (second ed.) Addison-Wesley (1968)

Michel, R.: On the influence of moments in nonuniform expansions of Chebyshev–Cramer type. Statistics and Decisions **1**, N. 2, 205–215 (1983)

Hipp, C.: Improved approximations for the aggregate claims distribution in the individual model. ASTIN Bull. **16**, 89–100 (1986)

Ibragimov, I.A.: On the accuracy of Gaussian approximation to the distribution functions of sums of independent variables. Theory Probab. Appl. **11**, 559–580 (1966)

Ibragimov, I.A.: On the Chebyshev–Cramer asymptotic expansions (In Russian). Teor. veroyatn. i primen., 506–519 (1967)

Ibragimov, I.A., and Linnik, Yu.V.: Independent and stationary sequences of random variables. Wolters-Noordhoff, Groningen (1971)

Kornya, P.: Distribution of aggregate claims in the Individual Risk Theory model. Society of Actuaries: Transactions. **35**, 823–858 (1983)

Kruopis, Yu.: Precision of approximations of the generalized Binomial distribution by convolutions of Poisson measures. Lithuanian Math. J. **26**, 37–49 (1986a)

Kruopis, Yu.: Approximations for distributions of sums of lattice random variables I. Lithuanian Math. J. **26**, 234–244 (1986b)

Le Cam, L.: An approximation theorem for the Poisson binomial distribution. Pacific J. Math. **10**, 1181–1197 (1960)

Nagaev, S.V., Chebotarev, V.I.: On asymptotic expansions of Bergström type in Hilbert space. (In Russian) Trudy Inst. Mat. Novosibirsk. **13**, 66–77 (1989)

Petrov, V.V.: Sums of Independent Random Variables. Springer-Verlag (1975)

Presman, E.L.: Approximation of binomial distributions by infinitely divisible ones. Theory Prob. Appl. **28**, 393–403 (1983)

Quine, M.P.: Probability approximations for divisible discrete distributions. Austral. J. Statist. **36**, 339–349 (1994)

Rachev, S.T. and Rüschendorf, L.: Approximation of sums by compound Poisson distributions with respect to stop-loss distances. Adv. Appl. Prob. **22**, 350–374 (1990)

Rozovski, L.V.: The accuracy of an estimate of the remainder term in the central limit theorem. Theor. Probab. Appl. **23**, 712–730 (1978)

Šiaulys, J., Čekanavičius, V.: Approximation of distributions of integer-valued additive functions by discrete charges. I. Lithuanian Math. J. **28**, 392–401 (1988)

Šiaulys, J., Čekanavičius, V.: Approximation of distributions of integer-valued additive functions by discrete charges. II. Lithuanian Math. J. **29**, 80–95 (1989)

Tsaregradskii, I.P.: On uniform approximation of a binomial distribution by infinitely divisible laws. (In Russian) Teorya Veroyatn. Primen. **3**, 470–474 (1958)