stochastic processes and their applications

www.elsevier.com/locate/spa

Signed Poisson approximations for Markov chains

V. Čekanavičius *, M. Mikalauskas

Department of Mathematics, Vilnius University, Naugarduko 24, 2006 Vilnius, Lithuania

Received 3 July 1997; received in revised form 1 June 1998; accepted 19 February 1999

Abstract

Consider a sum of Markov dependent lattice variables. The normal approximation is trivial for this sum if the total variation distance is considered. Replacement of the normal approximation by its Poisson structured analogue changes the situation radically. Moreover, considering the Markov binomial distribution we prove that signed Poisson approximation can be more accurate than both the normal and Poisson approximations. Possible improvements due to asymptotic expansions are discussed.

1991 MSC: 60J10; 60F05

Keywords: Signed Poisson approximation; Compound Poisson law; Total variation norm; Markov chain

1. Introduction

The aim of this paper is giving some possible alternatives to the normal law for sums of dependent lattice random variables. Applying normal approximation to lattice distributions we encounter two problems related to the differences in supports. First, the Edgeworth expansion contains additional summands to compensate the jumps of the approximated distribution. Second, it is impossible to get any but trivial estimate for all Borel sets. In general, other standard infinitely divisible approximations cannot improve the situation. For example, the standard Poisson approximation (even with long asymptotic expansions) is inapplicable for a sequence of random variables. In this paper we show that problems mentioned above can be solved using the signed Poisson approach, i.e., by replacing the normal approximation by its lattice Poisson-like analogue. Moreover, the signed Poisson approach allows us to use one approximation, instead of two or more. In Section 3 we prove that, for some parameters of the Markov binomial distribution, the signed Poisson approximation is of the same or better accuracy than the better one of the normal and Poisson approximations.

* Corresponding author.

E-mail address: vydas.cekanavicius@maf.vu.lt (V. Čekanavičius)

0304-4149/99/$ - see front matter
PII: S0304-4149(99)00017-4

For more detailed formulations we need the following notation. Let E_a denote the distribution concentrated at a point $a \in \mathbb{R}$, $E \equiv E_0$. All products and powers of measures are understood in the convolution sense, i.e., $FG\{A\} = \int_{\mathbb{R}} F\{A-x\}\, G\{\mathrm{d}x\}$, $F^0 \equiv E$. For a finite measure W we denote by $\widehat{W}(t)$ its Fourier–Stieltjes transform and by $\|W\|$ its total variation norm. We also denote $\exp\{W\} = \sum_{k=0}^{\infty} W^k/k!$ (the exponential of W) and $|W| = \sup_x |W\{(-\infty, x)\}|$ (an analogue of the uniform Kolmogorov distance). We denote by C positive absolute constants that can vary from line to line. Similarly, by $C(\cdot)$ we denote constants depending on the indicated arguments only. Sometimes, to avoid a possible confusion, we supply constants C with indices (as, e.g. in Eq. (1.3)). We always use the letter θ to denote a quantity satisfying $|\theta| \leqslant 1$. If F and G are finite measures, then

$$\|FG\| \leqslant \|F\|\,\|G\|, \quad \|\exp\{F\}\| \leqslant \exp\{\|F\|\}, \quad |F| \leqslant \|F\|,$$
$$\|F\|/2 \leqslant \sup_B |F\{B\}| \leqslant \|F\|, \tag{1.1}$$
$$\widehat{\exp\{F\}}(t) = \exp\{\widehat{F}(t)\}, \quad \widehat{FG}(t) = \widehat{F}(t)\widehat{G}(t), \quad \widehat{E}_a(t) = \mathrm{e}^{ita}, \quad \widehat{E}(t) = 1.$$

The supremum in Eq. (1.1) is taken over all Borel sets B. Note also that any Poisson distribution can be written in the form $\exp\{\lambda(E_1 - E)\}$ with $\lambda > 0$. In Sections 3 and 4 we also use the usual convention

$$\binom{a}{j} = \frac{a(a-1)\cdots(a-j+1)}{1\cdot 2\cdots j}. \tag{1.2}$$

For $a \in \mathbb{R}$ and a positive integer j. The real part of a complex number H is denoted by $Re\,H$.

Definition. Let $\lambda \in \mathbb{R}$, and let F be a distribution. Then $\exp\{\lambda(F - E)\}$ is called a signed compound Poisson measure. In particular, $\exp\{\lambda(E_1 - E)\}$ is called a signed Poisson measure.

Emphasize that, in comparison with a standard compound Poisson distribution we have for $\lambda < 0$, *signed* measures of finite variation. We have to note, however, that one can also find in the literature the terms *generalized Poisson* and *pseudo-Poisson* measures. We abbreviate convolutions of signed compound Poisson and signed Poisson measures by SCP and SP, respectively. In this paper, SP approximations are, as a rule, convolutions of two Poisson-like measures, while we use the abbreviation SCP when the compound measure has a more complicated structure. Note that all SCP measures are infinitely divisible.

It seems that Presman (1983) and Kornya (1983) were the first, who used SP measures as approximations (though other properties of the signed exponential measures were already used before, see Cuppens (1975)). The result of Presman (1983) is the following.

Theorem 1.1. *Let* $0 \leqslant p \leqslant 1/2$. *Then*

$$\|((1-p)E + pE_1)^n - \exp\{np(E_1 - E) - np^2(E_1 - E)^2/2\}\|$$
$$\leqslant C_1 \min(np^3, p^{3/2} n^{-1/2}). \tag{1.3}$$

Presman's approximation is an example of asymptotic expansion *in the exponent*. Remarkably, the standard Poisson approximation with one term of asymptotic expansion provides the rate of accuracy $\min(np^3, p^2)$, (see, e.g., Barbour et al. (1992)) which is always greater than (or, at least, equal) to the estimate in the right-hand side of (1.3). Thus, asymptotics in the exponent favourably differ from the standard asymptotics. On the other hand, one can consider Presman's approximation as a discrete analogue of the normal law (it ensures the matching of two moments). The Berry–Esseen, theorem, however, provides the rate of accuracy $(np)^{-1/2}$, which is always greater than (or, at least, equal) to the right-hand side of Eq. (1.3). Thus, we conclude that the SP approach can produce approximations essentially better than the known asymptotic expansions. One easily checks that Presman's approximation can be also written as $\exp\{(np - np^2)(E_1 - E) - np^2(E_2 - E)/2\}$.

For independent summands, SP and SCP approximations were successfully applied in many fields, see Hipp (1986), Kruopis (1986a,b), Borovkov (1988), Čekanavičius (1996, 1997, 1998) and references therein. On the other hand, it seems that only Borovkov and Pfeifer (1996) considered SCP approximations for the distributions of Markov chains. We do not review, however, their result, since our paper is devoted to a scheme different from theirs. We consider *sums* of Markov dependent variables. They are constantly getting a lot of attention, see Dobrushin (1953), Serfling (1975), Sirazhdinov and Formanov (1979), Wang (1981, 1992), Gani (1982), Serfozo (1986), Barbour et al. (1992, Section 8.5), Geske et al. (1995) and references therein. The authors mentioned above considered the normal, Poisson, and compound Poisson approximations. Thus, in our paper we obtain a refinement of some known results using a new signed Poisson approach. We also show how asymptotic expansions to the known Poisson or compound Poisson approximations can be constructed.

The structure of this paper is the following. In Section 2, we consider a sequence of Markov dependent variables and show benefits of the replacement of normal approximation by its SP analogue. In Sections 3 and 4, we use the sum of Markov dependent Bernoulli variables for proving the universality of SP and SCP approximations. In Section 5, we obtain some local estimates. Section 6 contains concluding remarks.

2. SP approximation for a sequence of Markov dependent variables

Let $\xi_0, \xi_1, \ldots, \xi_n, \ldots,$ be an s-state Markov chain with the transition matrix $P = (p_{ij})_1^s$ and initial distribution Π. Suppose P is irreducible and aperiodic. Note that we deal with a sequence of variables. Therefore ξ_j and P do not depend on n. The main purpose of this Section is the demonstration of a greater flexibility of SP measure, in comparison with the standard Poisson law. As a rule, the standard Poisson approximation is used in the scheme of series, and not for the sequence of random variables. We also show that, in comparison with the normal law, the SP approximation also has at least two advantages. Both are related to the fact that an approximated measure and its approximation are concentrated on the same lattice. Thus, we can use the total variation distance and avoid additional summands in asymptotic expansions, which are necessary for the normal approximation. Moreover, as far as both approximations are on the same lattice, the estimates do not depend on the maximum span of the lattice.

There are many results on the normal approximation to Markov dependent variables. We chose a quite partial situation allowing us to demonstrate some benefits of the SP approach more clearly. We use the approach of Sirazhdinov and Formanov (1979). Let f be a bounded function. Set

$$S_{n0} = \sum_{j=0}^{n} f(\xi_j). \tag{2.1}$$

Let $h_j = f(j)$, $j = 1,\dots,s$. Denote by $\eta_j(n)$ the time spent in state j until the moment n, i.e.,

$$\eta_j(n) = \sum_{k=0}^{n} I\{\xi_k = j\},$$

where $I\{\cdot\}$ denotes the indicator function. Then

$$S_{n0} = \sum_{j=0}^{s} h_j \eta_j(n). \tag{2.2}$$

We denote by F_{n0} the distribution of S_{n0}. Let $\lambda_M(t_1,t_2,\dots,t_s)$ denote the eigenvalue of $Q(t_1,t_2,\dots,t_s) = (p_{kj}\exp\{it_j\})_1^s$ having the greatest absolute value. Set

$$p_j = \frac{1}{\mathrm{i}}\,\frac{\partial \ln \lambda_M}{\partial t_j}\bigg|_{t_1=\cdots=t_s=0}, \qquad \lambda_{kj} = -\frac{\partial^2 \ln \lambda_M}{\partial t_k \partial t_j}\bigg|_{t_1=\cdots=t_s=0}.$$

Sirazhdinov and Formanov (1979, p.14) showed that, in a neighborhood of zero, the characteristic function of $(\eta_1(n),\eta_2(n),\dots,\eta_s(n))$ can be expressed as the sum

$$\lambda_M^n(t_1,\dots,t_s)A(t_1,\dots,t_s) + \theta C(P,s)|t|\rho_1^n, \quad 0 < \rho_1 < 1. \tag{2.3}$$

Here

$$A(t_1,t_2,\dots,t_s) = \Pi^{\mathrm{T}}(t_1,t_2,\dots,t_s)\,x\,(t_1,t_2,\dots,t_s)y^{\mathrm{T}}(t_1,t_2,\dots,t_s)\,\mathbf{e},$$

$$\mathbf{e} = (1,1,\dots,1)^{\mathrm{T}}, \quad \Pi^{\mathrm{T}}(t_1,t_2,\dots,t_s) = (\Pi_1\mathrm{e}^{\mathrm{i}t_1},\Pi_2\mathrm{e}^{\mathrm{i}t_2},\dots,\Pi_s\mathrm{e}^{\mathrm{i}t_s}),$$

$$(\Pi_1,\Pi_2,\dots,\Pi_s) = \Pi,$$

and $x(t_1,t_2,\dots,t_s)$ and $y(t_1,t_2,\dots,t_s)$ are the eigenvectors of $Q(t_1,t_2,\dots,t_s)$ corresponding to $\lambda_M(t_1,t_2,\dots,t_s)$:

$$y^{\mathrm{T}}(t_1,t_2,\dots,t_s)\,Q\,(t_1,t_2,\dots,t_s) = \lambda_M(t_1,t_2,\dots,t_s)y^{\mathrm{T}}(t_1,t_2,\dots,t_s),$$

$$Q(t_1,t_2,\dots,t_s)\,x\,(t_1,t_2,\dots,t_s) = \lambda_M(t_1,t_2,\dots,t_s)\,x\,(t_1,t_2,\dots,t_s),$$

$$y^{\mathrm{T}}(t_1,t_2,\dots,t_s)\,x\,(t_1,t_2,\dots,t_s) = 1.$$

$C(P,s)$ is a non-negative constant depending on P and s. Note that, in Sirazhdinov (1952), $A(t_1,t_2,\dots,t_s)$ is expressed in terms of determinant of $Q(t_1,t_2,\dots,t_s)$.

Let $\mu_1, \mu_2, \ldots, \mu_s$ be the numbers defined by the formal equation

$$\mu_j = \frac{\partial A(t_1, t_2, \ldots, t_s)}{\partial t_j}\bigg|_{t_1=\cdots=t_s=0}. \tag{2.4}$$

The numbers $\mu_1, \ldots, \mu_s$ depend on the initial distribution Π. Set

$$K = \sum_{j=1}^{s} \mu_j h_j, \qquad L = \left(\sum_{j=1}^{s} h_j \frac{\partial}{\partial t_j}\right)^3 \ln \lambda_M(t_1, \ldots, t_s)\,|_{t_1=\cdots=t_s=0}\,,$$

$$n\mu = n \sum_{j=1}^{s} h_j p_j, \qquad n\sigma^2 = n \sum_{k,j} \lambda_{kj} h_k h_j. \tag{2.5}$$

If $\sigma^2 > 0$, then we can define the standardized sum

$$S_{ns} = (S_{n0} - n\mu)/(\sigma\sqrt{n}).$$

Let F_{ns} be the distribution of S_{ns}, and let Φ be the standard normal distribution. Set

$$M(t) = \frac{\mathrm{i}t}{\sigma\sqrt{n}} K + \frac{(\mathrm{i}t)^3}{6\sigma^3\sqrt{n}} L.$$

Suppose $\Phi_1(x)$ has the Fourier–Stieltjes transform $\mathrm{e}^{-t^2/2} M(t)$. Let $Z(x) = \mathscr{S}(\sigma\sqrt{n}x + n\mu)\mathrm{e}^{-x^2/2}(2\pi n\sigma^2)^{-1/2}$, where $\mathscr{S}(x) = [x] - x + 1/2$ and $[x]$ denotes the integer part of x.

Now we can introduce the main assumptions of this section. We assume that

$$\text{among } h_1, \ldots, h_s, \text{ there are at least two different numbers,} \tag{2.6}$$

$$h_1, h_2, \ldots, h_s \text{ are integers with the gr.c.d. equal to 1,} \tag{2.7}$$

$$\text{the quadratic form } \sum_{k,j=1}^{s-1} \lambda_{kj} t_k t_j \text{ is positive definite.} \tag{2.8}$$

Note that Eq. (2.8) implies $\sigma^2 > 0$. Assumption (2.6) is not very restrictive, since in the case of coinciding h_j we have $S_{n0} \equiv \text{Const}$. It is possible to replace Eq. (2.7) by the requirement that all h_j belong to the same lattice. However, a suitable centering and norming would produce the same situation as in Eq. (2.7), see, e.g., Gnedenko and Kolmogorov (1954, p. 232).

Theorem 2.1 follows immediately from the non-uniform estimates obtained in Sirazhdinov and Formanov (1979, p. 36).

Theorem 2.1. *Let assumptions* (2.6)–(2.8) *be satisfied. Then*

$$\sup_x |F_{ns}(x) - \Phi(x) - \Phi_1(x) - Z(x)| = \mathrm{o}(n^{-1/2}). \tag{2.9}$$

We shall replace Φ by its SP analogue. Set

$$G_0 = \exp\left\{\frac{\sigma^2 + \mu}{2}(E_1 - E) + \frac{\sigma^2 - \mu}{2}(E_{-1} - E)\right\}, \tag{2.10}$$

$$G_{01} = K(E_1 - E) + n(L - \mu)(E_1 - E)^3/6, \tag{2.11}$$

$$\widehat{H}_0(t) = \widehat{G}_0(t/(\sigma\sqrt{n}))\exp\{-\mathrm{i}t\mu/(\sigma\sqrt{n})\}, \qquad \widehat{H}_{01}(t) = \widehat{G}_{01}(t/(\sigma\sqrt{n})), \tag{2.12}$$

$$\widehat{\Delta}_0(t) = \widehat{F}_{ns}(t) - \widehat{H}_0^n(t)(1+\widehat{H}_{01}(t)), \qquad \widehat{\Delta}_{01}(t) = \widehat{F}_{ns}(t) - \mathrm{e}^{-t^2/2}(1+M(t)),$$
$$\widehat{\Delta}_{02}(t) = \mathrm{e}^{-t^2/2}(1+M(t)) - \widehat{H}_0^n(t)(1+\widehat{H}_{01}(t)). \tag{2.13}$$

We can now formulate the main result of this section.

Theorem 2.2. *Let assumptions* (2.6)–(2.8) *be satisfied. Then*

$$\|F_{n0} - G_0^n\| = \mathrm{O}(n^{-1/2}), \tag{2.14}$$

$$\|F_{n0} - G_0^n(E+G_{01})\| = \mathrm{o}(n^{-1/2}). \tag{2.15}$$

Remark 2.1. The measure G_0 was introduced in Kruopis (1986b) who proposed to call it the normal–Poisson approximation.

Remark 2.2. For the sums of independent random variables, some analogues of Eqs. (2.14)–(2.15) were considered in Čekanavičius (1998).

Remark 2.3. Though one can center and norm both F_{n0} and G_0 as in Theorem 2.1, this hardly makes sense, since both measures are concentrated on the same lattice.

Note that G_0 satisfies simple recursions and, hence, can be used in practical calculations, see Kruopis (1986b).

Proof of Theorem 2.2. We will prove Eq. (2.15) only. Estimate (2.14) can be obtained similarly. Since F_{n0}, G_0, and G_{01} are concentrated on integers and $\sigma < \infty$, we can use the approach of Presman (1983). We omit t for brevity. By the inversion formula of Presman (1983) we have

$$\begin{aligned}\|F_{n0} - G_0^n(E+G_{01})\| &\leqslant C\int_{|t|\leqslant\pi\sigma\sqrt{n}} (|\widehat{\Delta}_0| + |\widehat{\Delta}_0''|)\mathrm{d}t \\ &\leqslant C\int_{\delta n^{1/6}\leqslant|t|\leqslant\delta n^{1/2}} + C\int_{\delta n^{1/2}\leqslant|t|\leqslant\pi\sigma n^{1/2}} \\ &\quad + C\int_{|t|\leqslant\delta n^{1/6}} = J_1 + J_2 + J_3,\end{aligned} \tag{2.16}$$

$$J_3 \leqslant C\int_{|t|\leqslant\delta n^{1/6}} (|\widehat{\Delta}_{01}| + |\widehat{\Delta}_{01}''|)\mathrm{d}t + \int_{|t|\leqslant\delta n^{1/6}} (|\widehat{\Delta}_{02}| + |\widehat{\Delta}_{02}''|)\mathrm{d}t = J_{31} + J_{32} \tag{2.17}$$

for a positive number $\delta > 0$. We further need some auxiliary results from Sirazhdinov and Formanov (1979, pp. 31–33).

Lemma 2.1. *Let assumptions* (2.6)–(2.8) *be satisfied. Then there exist numbers* δ, $\delta_2 > 0$, $0 < \rho < 1$, *and a function* $\delta_1(n)$ *such that, for* $|t| \leqslant \delta n^{1/6}$ *and* $j \geqslant 0$,

$$\left|\frac{\mathrm{d}^j \widehat{\Delta}_{01}}{\mathrm{d}t^j}\right| \leqslant \frac{\delta_1(n)}{\sqrt{n}}(\max(1,|t|))^{\max(2j,j+3)}\mathrm{e}^{-\delta_2 t^2} + \mathrm{O}(\rho^n) \tag{2.18}$$

and $\lim_n \delta_1(n) = 0$. *Furthermore, for all* $|t| \leqslant \delta n^{1/2}$,

$$\left|\frac{\mathrm{d}^j \widehat{F}_{ns}}{\mathrm{d}t^j}\right| \leqslant C(j,s,P)\max(1,|t|)\mathrm{e}^{-\delta_2 t^2}, \tag{2.19}$$

where $C(j,s,P)$ *is a constant depending on* j,s *and* P.

Lemma 2.2. *Let assumptions* (2.6)–(2.8) *be satisfied. Then, for* $j \geqslant 0$ *and* $0 < \varepsilon < |t| < \pi$,

$$|\widehat{F}_{ns}^{(j)}(\sigma\sqrt{n}t) = \mathrm{O}(\rho^n), \quad 0 < \rho < 1. \tag{2.20}$$

By the definition of G_0 we have, for all $|t| \leqslant \pi$,

$$|\widehat{G}_0(t)| \leqslant \exp\{-\sigma^2 \sin^2(t/2)\} \leqslant \exp\{-\sigma^2 t^2/\pi^2\}.$$

Hence

$$|\widehat{H}_0^n(t)| \leqslant \mathrm{e}^{-Ct^2}. \tag{2.21}$$

Combining (2.16)–(2.19) we obtain

$$J_1 + J_2 + J_{31} = \mathrm{o}(n^{-1/2}). \tag{2.22}$$

It remains to estimate J_{32}. We have

$$\begin{aligned} |\widehat{\Delta}_{02}| \leqslant\ & |\widehat{H}_0^n - \mathrm{e}^{-t^2/2} - n\mathrm{e}^{-t^2/2}(\widehat{H}_0 \mathrm{e}^{t^2/(2n)} - 1)|\,|1 + \widehat{H}_{01}| \\ & + n\mathrm{e}^{-t^2/2}|\,\widehat{H}_0 \mathrm{e}^{t^2/(2n)} - 1 - \mu(\mathrm{i}t)^3/(6\sigma^3 n^{3/2})\,|\,|1 + \widehat{H}_{01}| \\ & + \mathrm{e}^{-t^2/2}|(1 + \mu(\mathrm{i}t)^3/(6\sigma^3 n^{1/2}))(1 + \widehat{H}_{01}) - (1 + M(t))|. \end{aligned} \tag{2.23}$$

Applying Bergström's (1951) identity

$$a^n - b^n - nb^{n-1}(a-b) = \sum_{j=2}^{n}(j-1)a^{n-j}b^{j-2}(a-b)^2 \tag{2.24}$$

and the trivial estimate $|t|^k \exp\{-Ct^2\} \leqslant C(k)\exp\{-Ct^2/2\}$, we obtain

$$|\widehat{\Delta}_{02}| \leqslant Cn^{-1}\mathrm{e}^{-Ct^2}. \tag{2.25}$$

It is not difficult to verify an analogous estimate for $\widehat{\Delta}_{02}''$:

$$|\widehat{\Delta}_{02}''| \leqslant Cn^{-1}\mathrm{e}^{-Ct^2}. \tag{2.26}$$

From estimates (2.25), (2.26) and (2.16), (2.22) we get the statement of the theorem. □

3. Approximation of the Markov binomial distribution

In this section, we obtain an analogue of Eq. (1.3), thus proving the universality of SP approximation for a sum of Markov dependent Bernoulli variables. Unlike the previous section, now we consider the scheme of series, when the transition probabilities may depend on n.

The Markov binomial distribution was studied by many authors. Among numerous publications devoted to this case, we would like to refer the reader, for example, to Koopman (1950), Dobrushin (1953), Serfling (1975), Wang (1981, 1992), Gani (1982), Serfozo (1986), and references therein. These authors usually considered the Poisson or compound Poisson approximations only. Some estimates of the accuracy of approximations were established. However, with few exceptions all these estimates depend on the existence of limiting Poisson (compound Poisson) law. But such an existence means that parameters should be small. We do not study such partial cases preferring to concentrate on a larger set of parameters covering, simultaneously, a few possible limiting distributions. The main attention is paid to the rate of approximation. We must also note that the Markov binomial distribution can significantly differ from the binomial distribution. Thus, for example, it has at least seven limiting laws, some of which are not even infinitely divisible, see Dobrushin (1953).

Let $\xi_0, \xi_1, \xi_2, \ldots, \xi_n$ be a Markov chain with the initial distribution

$$P(\xi_0 = 1) = p_0, \qquad P(\xi_0 = 0) = 1 - p_0.$$

We assume that

$$P(\xi_i = 1 | \xi_{i-1} = 1) = p, \qquad P(\xi_i = 0 | \xi_{i-1} = 1) = q,$$

$$P(\xi_i = 1 | \xi_{i-1} = 0) = \bar{q}, \qquad P(\xi_i = 0 | \xi_{i-1} = 0) = \bar{p}, \qquad p + q = \bar{q} + \bar{p} = 1.$$

Let the transition matrix be

$$P = \begin{pmatrix} p & q \\ \bar{q} & \bar{p} \end{pmatrix}. \tag{3.1}$$

Set

$$S_n = \xi_1 + \xi_2 + \cdots + \xi_n.$$

Denote the distribution of S_n by F_n. We call F_n a Markov binomial distribution. Note that, in the literature, the definition of the Markov binomial distribution slightly varies from paper to paper (e.g., sometimes ξ_0 is added to the sum S_n). Our definition corresponds to that of Serfozo (1986). The compound Poisson limit occurs when $p_0 \to 0$, $n\bar{q} \to \lambda$, and $p \to \tilde{p} < 1$ as $n \to \infty$. If $\tilde{p} = 0$, then we have the Poisson limit. Therefore, the natural assumption should be $\bar{q} + p \to 0$. But in this case we would not cover the classical normal case, where $\bar{q}$ and p are constants. Therefore, we assume that $\bar{q}$ and p are small enough (the "smallness" is determined by the method of proof), but not necessarily vanishing. Let

$$p \leqslant 1/20, \qquad \bar{q}/(q + \bar{q}) \leqslant 1/30. \tag{3.2}$$

Note that, under Eq. (3.2), the limiting distribution can be the compound Poisson (including Poisson), degenerate, or normal distribution.

We further need some notation. Set

$$v_1 = \frac{\bar{q}}{q+\bar{q}}, \qquad v_2 = \frac{2q\bar{q}(p-\bar{q})}{(q+\bar{q})^3}, \tag{3.3}$$

$$G_1 = \exp\{v_1(E_1 - E) + (v_2 - v_1^2)(E_1 - E)^2/2\}, \tag{3.4}$$

$$G_2 = \exp\{(v_1 + (v_2 - v_1^2)/2)(E_1 - E) + (v_2 - v_1^2)(E_{-1} - E)/2\}. \tag{3.5}$$

We further assume that $n \geqslant 1$. We can now formulate the main result of this section.

Theorem 3.1. *Let assumption* (3.2) *be satisfied. Then*

$$\|F_n - G_1^n\| \leqslant C_2(p+\bar{q})^2 \min((n\bar{q})^{-1/2}, n\bar{q}) + C_3|p-\bar{q}|\min(1,(n\bar{q})^{-1/2}) \tag{3.7}$$

and

$$\|F_n - G_2^n\| \leqslant C_4(p+\bar{q}) \min((n\bar{q})^{-1/2}, n\bar{q}) + C_3|p-\bar{q}|\min(1,(n\bar{q})^{-1/2}). \tag{3.8}$$

Before proving Theorem 3.1, we want to discuss some aspects of estimates (3.7)–(3.8). If $p=\bar{q}$, then S_n becomes the sum of independent Bernoulli variables. Consequently, F_n becomes the binomial distribution. Then (3.7) is equivalent to Presman's result (1.3). (However, our assumption (3.2) is stronger than merely $p \leqslant 1/2$.) Now consider the case $p, \bar{q} \sim$ const. Both estimates (3.7) and (3.8) are of order $n^{-1/2}$, i.e., both approximations are then comparable to the normal one (and hold for a stronger metric). Both G_1 and G_2 are also comparable with the Poisson approximation. Very sharp results were obtained for $\bar{q} = o(n^{-1})$, but not for larger p and $\bar{q}$. For a stationary Markov chain, a sharp general result can be found in Barbour et al. (1992, p. 165), but our case is not necessarily stationary. Therefore, we state a result for the Poisson approximation.

Theorem 3.2. *Let assumption* (3.2) *be satisfied. Then*

$$\|F_n - \exp\{nv_1(E_1 - E)\}\| \leqslant C_5(p+\bar{q})\min(1, n\bar{q}) + C_3|\bar{q}-p|\min(1,(n\bar{q})^{-1/2}). \tag{3.9}$$

For $\bar{q} = p$, estimate (3.9) is of the right order and coincides with the result of Prokhorov (1953). Comparing Eqs. (3.7) and (3.8) with Eq. (3.9) we see that both SP approximations are sharper in the sense of order (or, at least, of the same order) than the Poisson approximation. To compare G_1 and G_2, first note that the approximation G_2 corresponds to the normal-Poisson case (recall Section 1). Moreover, v_1 and $v_2 - v_1^2$ are the main parts of $\mathbf{E}S_n/n$ and $\mathbf{Var}\, S_n/n - \mathbf{E}S_n/n$ respectively. On the other hand, as one can notice from Eqs. (3.7) and (3.8), G_1 is an analogue of Presman's (1.3) approximation and is sharper than G_2 for small values of parameters.

SP approximations can be improved by asymptotic expansions. We consider one example only. Set

$$v_3 = 6q\bar{q}(\bar{q}-p)(\bar{q}+q(\bar{q}-p))/(q+\bar{q})^5, \tag{3.10}$$

$$A_1 = (\bar{q} - p)(\bar{q} - p_0(q + \bar{q}))/(q + \bar{q})^2, \qquad A_2 = \nu_3/6 - \nu_1(\nu_2 - \nu_1^2)/2 - \nu_1^3/3. \tag{3.11}$$

Theorem 3.3. *Let assumption* (3.2) *be satisfied. Then*

$$||F_n - G_1^n\{E + A_1(E_1 - E) + A_2 n(E_1 - E)^3\}||$$
$$\leqslant C_6(p + \bar{q})^3 \min((n\bar{q})^{-1}, n\bar{q}) + C_7|\bar{q} - p|(\bar{q} + p)\min((n\bar{q})^{-1}, 1). \tag{3.12}$$

Remark 3.1. If p and $\bar{q}$ are constants, then Eq. (3.12) is of order $O(n^{-1})$. Thus Eq. (3.12) is an analogue of the Edgeworth expansion. It is specific that it holds for all Borel sets and needs no additional $\mathscr{S}$-like summands, just like the SP expansions considered in Section 2.

Proof of Theorem 3.1. The proof is based on Perron's formula. Solving the characteristic equation we obtain

$$\widehat{F}_n(t) = \widehat{\Lambda}_1^n(t)\widehat{W}_1(t) + \widehat{\Lambda}_2^n(t)\widehat{W}_2(t), \qquad \widehat{\Lambda}_{1,2}(t) = (p\mathrm{e}^{\mathrm{i}t} + \bar{p} \pm \widehat{D}^{1/2}(t))/2,$$

$$\widehat{W}_{1,2}(t) = \frac{p_0}{2}(1 \pm (q + \bar{q} + p(\mathrm{e}^{\mathrm{i}t} - 1))\widehat{D}^{-1/2}(t))$$
$$+ \frac{(1 - p_0)}{2}(1 \pm (q + \bar{q} + p(\mathrm{e}^{\mathrm{i}t} - 1) + 2(\bar{q} - p)(\mathrm{e}^{\mathrm{i}t} - 1))\widehat{D}^{-1/2}(t)),$$

where $\widehat{D}(t)$ denotes the discriminant of the characteristic equation. Note that, by Eq. (3.2), it can be expressed in the following way:

$$\widehat{D}(t) = (p\mathrm{e}^{\mathrm{i}t} + \bar{p})^2 + 4\mathrm{e}^{\mathrm{i}t}(\bar{q} - p) = (1 + \bar{q} - p\mathrm{e}^{\mathrm{i}t})^2(1 + 4\bar{q}(\mathrm{e}^{\mathrm{i}t} - 1)/(1 + \bar{q} - p\mathrm{e}^{\mathrm{i}t})^2)$$
$$= (q + \bar{q} + p(\mathrm{e}^{\mathrm{i}t} - 1))^2(1 + 4q(\bar{q} - p)(\mathrm{e}^{\mathrm{i}t} - 1)/(q + \bar{q} + p(\mathrm{e}^{\mathrm{i}t} - 1))^2). \tag{3.13}$$

We further introduce some auxiliary finite measures. Set

$$B = \frac{1}{q + \bar{q}} \sum_{j=0}^{\infty} \left(\frac{p}{q + \bar{q}}\right)^j (E_1 - E)^j, \qquad Y = 4\bar{q}(E_1 - E)B^2, \tag{3.14}$$

$$\widetilde{B} = \frac{1}{q + \bar{q}} \sum_{j=0}^{\infty} \left(\frac{-p}{q + \bar{q}}\right)^j (E_1 - E)^j, \qquad \widetilde{Y} = 4q(\bar{q} - p)(E_1 - E)\widetilde{B}^2,$$

$$\widetilde{D} = \sum_{j=0}^{\infty} \binom{-1/2}{j} \widetilde{Y}^j. \tag{3.15}$$

Let Λ_1 and Λ_2 be finite measures corresponding to the eigenvalues $\widehat{\Lambda}_1(t)$ and $\widehat{\Lambda}_2(t)$, respectively. Then

$$\Lambda_1 = \frac{1}{2}\left(pE_1 + \bar{p}E + ((1 + \bar{q})E - pE_1)\sum_{j=0}^{\infty} \binom{1/2}{j} Y^j\right), \tag{3.16}$$

$$\Lambda_2 = \frac{1}{2}\left(pE_1 + \bar{p}E - ((1+\bar{q})E - pE_1)\sum_{j=0}^{\infty}\binom{1/2}{j} Y^j \right), \tag{3.17}$$

$$F_n = \Lambda_1^n W_1 + \Lambda_2^n W_2, \tag{3.18}$$

$$\begin{aligned} W_1 &= \frac{1}{2}\{p_0\{E + ((q+\bar{q})E + p(E_1 - E))\widetilde{B}\widetilde{D}\} \\ &\quad + (1-p_0)\{E + ((q+\bar{q})E + p(E_1 - E) + 2(\bar{q}-p)(E_1 - E))\widetilde{B}\widetilde{D}\} \\ &= \frac{1}{2}\{E + \widetilde{D}\} + (1-p_0)(\bar{q}-p)(E_1 - E)\widetilde{B}\widetilde{D}, \end{aligned} \tag{3.19}$$

$$W_2 = \frac{1}{2}\{E - \widetilde{D}\} - (1-p_0)(\bar{q}-p)(E_1 - E)\widetilde{B}\widetilde{D}. \tag{3.20}$$

(Note that $((q+\bar{q})E + p(E_1 - E))\widetilde{B} \equiv E$.)

We intentionally used two different expansions of the discriminant for Λ_i and W_i. By the properties of total variation norm we have

$$\begin{aligned} ||F_n - G_i^n|| &\leqslant ||\Lambda_1^n W_1 - G_i^n|| + ||\Lambda_2||^n ||W_2|| \\ &\leqslant ||\Lambda_1^n - G_i^n||\,||W_1|| + ||G_i^n(W_1 - E)|| + ||\Lambda_2||^n ||W_2||. \end{aligned} \tag{3.21}$$

By Eq. (3.2) we have $||W_1|| \leqslant C$, $||B|| \leqslant 1/(q+\bar{q}) \leqslant 20/19$, $||Y|| \leqslant 1/3$, $||\widetilde{B}|| \leqslant 20/17$ and $||\widetilde{Y}|| \leqslant 2/3$. Hence

$$||W_2|| \leqslant (1-p_0)|\bar{q}-p|\,||E_1 - E||\,||\widetilde{B}||\,||\widetilde{D}|| + ||\widetilde{Y}||\sum_{j=1}^{\infty}||\widetilde{Y}||^{j-1}/2 \leqslant C|\bar{q}-p|. \tag{3.22}$$

Quite similarly we establish that, for $i = 1, 2$,

$$||G_i^n(W_1 - E)|| \leqslant ||G_i^n(E_1 - E)||C|\bar{q}-p|. \tag{3.23}$$

For the estimates with respect to the total variation distance we will use the following auxiliary result (see Presman (1985) or Šiaulys and Čekanavičius (1988)).

Lemma 3.1. *Let* $\widehat{R}(t) = \sum_{j=-\infty}^{\infty} R_j \exp\{itj\}, \sum_{j=-\infty}^{\infty}|R_j| < \infty$. *Then, for all* $\gamma > 0$, *a and* $v \in \mathbb{R}$,

$$\left(\sum_{j=-\infty}^{\infty}|R_j|\right)^2 \leqslant \left(\frac{1}{2} + \frac{1}{2\pi\gamma}\right)\int_{-\pi}^{\pi}\left(\gamma|\widehat{R}(t)|^2 + \frac{1}{\gamma}|(\widehat{R}(t)\mathrm{e}^{-\mathrm{i}tv})'|^2\right)\mathrm{d}t.$$

Applying Lemma 3.1 with $v = nv_1$ and $\gamma = \max(1, (n\bar{q})^{-1/2})$), one easily shows that the right-hand side of Eq. (3.23) is less than $C|\bar{q}-p|\min(1,(n\bar{q})^{-1/2})$.

By properties of the total variation norm we have

$$\begin{aligned} 2||\Lambda_2|| &\leqslant ||pE_1 + \bar{p}E - (1+\bar{q})E + pE_1|| \\ &\quad + ||(1+\bar{q})E - pE_1||\,||Y||\sum_{j=1}^{\infty}\left|\binom{1/2}{j}\right|\,||Y||^{j-1} \\ &\leqslant 1/2 + 2(p+\bar{q})(1+1/2) \leqslant 51/60. \end{aligned} \tag{3.24}$$

Thus

$$||\Lambda_2||^n \leqslant (51/120)^n \leqslant \mathrm{e}^{-n\ln 2} = \mathrm{e}^{-Cn}. \tag{3.25}$$

From Eqs. (3.21), (3.22), and (3.25) we see that to finish the proof it suffices to estimate $||\Lambda_1^n - G_i^n||$. Expanding $\widehat{\Lambda}_1(t)$ in powers of $\mathrm{e}^{\mathrm{i}t}-1$ we obtain

$$\widehat{\Lambda}_1(t) = 1 + v_1(\mathrm{e}^{\mathrm{i}t}-1) + \frac{v_2}{2}(\mathrm{e}^{\mathrm{i}t}-1)^2 + \frac{v_3}{6}(\mathrm{e}^{\mathrm{i}t}-1)^3 + C\theta\bar{q}(\bar{q}+p)^3|\mathrm{e}^{\mathrm{i}t}-1|^4, \tag{3.26}$$

$$\widehat{\Lambda}_1'(t) = \mathrm{i}\mathrm{e}^{\mathrm{i}t}\left(v_1 + v_2(\mathrm{e}^{\mathrm{i}t}-1) + \frac{v_3}{2}(\mathrm{e}^{\mathrm{i}t}-1)^2\right) + C\theta\bar{q}(\bar{q}+p)^3|\mathrm{e}^{\mathrm{i}t}-1|^3. \tag{3.27}$$

Taking into account the form of G_i after quite standard calculations we get

$$|\widehat{\Lambda}_1(t) - \widehat{G}_1(t)| \leqslant C\bar{q}(p+\bar{q})^2|t|^3, \qquad |\widehat{\Lambda}_1'(t) - \widehat{G}_1'(t)| \leqslant C\bar{q}(p+\bar{q})^2t^2, \tag{3.28}$$

$$|(\mathrm{e}^{-\mathrm{i}tv_1}\Lambda_1(t))'| \leqslant C\bar{q}|t|, \qquad |(\mathrm{e}^{-\mathrm{i}tv_1}\widehat{G}_j(t))'| \leqslant C\bar{q}|t| \quad (j=1,2), \tag{3.29}$$

$$|\widehat{\Lambda}_1(t) - \widehat{G}_2(t)| \leqslant C\bar{q}(p+\bar{q})|t|^3, \qquad |\widehat{\Lambda}_1'(t) - \widehat{G}_2'(t)| \leqslant C\bar{q}(p+\bar{q})t^2. \tag{3.30}$$

Moreover, we have, for $|t| \leqslant \pi$,

$$\begin{aligned}
|\widehat{\Lambda}_1(t)| &\leqslant \frac{1}{2}|p\mathrm{e}^{\mathrm{i}t} + \bar{p} + (1+\bar{q}-p\mathrm{e}^{\mathrm{i}t})(1+\widehat{Y}(t)/2)| + \frac{3}{8}|1+\bar{q}-p\mathrm{e}^{\mathrm{i}t}|\,|\widehat{Y}(t)|^2 \\
&\leqslant |1 + v_1(\mathrm{e}^{\mathrm{i}t}-1) + \bar{q}p|\mathrm{e}^{\mathrm{i}t}-1|^2(q+\bar{q})^2(1-2p/(q+\bar{q}))^{-1} \\
&\quad + 3\bar{q}^2|\mathrm{e}^{\mathrm{i}t}-1|^2(q+\bar{q})^{-3}/2 \\
&\leqslant 1 - 2v_1(1-v_1)\sin^2(t/2) + 4v_1\sin^2(t/2)/17 + v_1^2\sin^2(t/2)120/19 \\
&\leqslant 1 - 2v_1\sin^2(t/2)(1 - v_1 - 2/17 - 60v_1/19) \\
&\leqslant 1 - 24v_1\sin^2(t/2)/17 \leqslant \exp\{-Ct^2\bar{q}\}.
\end{aligned} \tag{3.31}$$

It is easy to check that, for $|t| \leqslant \pi$, the analogous estimates hold for $\widehat{G}_1(t)$ and $\widehat{G}_2(t)$:

$$|\widehat{G}_1(t)| \leqslant \mathrm{e}^{-Ct^2\bar{q}}, \qquad |\widehat{G}_2(t)| \leqslant \mathrm{e}^{-Ct^2\bar{q}}. \tag{3.32}$$

Combining Eqs. (3.28)–(3.32) and applying Lemma 3.1 with $\gamma = \max(1, \sqrt{n})$ and $v = nv_1$ we obtain

$$||\Lambda_1^n - G_1^n|| \leqslant C(p+\bar{q})^2\min((n\bar{q})^{-1/2}, n\bar{q}), \tag{3.33}$$

$$||\Lambda_1^n - G_2^n|| \leqslant C(p+\bar{q})\min((n\bar{q})^{-1/2}, n\bar{q}). \tag{3.34}$$

The statement of the theorem now follows from Eqs. (3.21)–(3.23), (3.25), (3.33) and (3.34). □

Proof of Theorem 3.2. The proof is very similar to that of Theorem 3.1. The only exception is that Eqs. (3.28)–(3.29) are replaced by

$$|\widehat{\Lambda}_1(t) - \widehat{P}_1(t)| \leqslant C\bar{q}(p+\bar{q})t^2, \qquad |\widehat{\Lambda}_1'(t) - \widehat{P}_1'(t)| \leqslant C\bar{q}(p+\bar{q})|t|,$$

$$|(\mathrm{e}^{-\mathrm{i}tv_1}\widehat{P}_1(t))'| \leqslant C\bar{q}|t|.$$

where $\widehat{P}_1(t) = \exp\{v_1(\mathrm{e}^{\mathrm{i}t}-1)\}$. □

Proof of Theorem 3.3. By the properties of variation norm, Bergström's identity (2.24), and Eqs. (3.14)–(3.20) we get

$$\begin{aligned}
&||F_n - G_1^n(E + A_1(E_1 - E) + nA_2(E_1 - E)^3||\\
&\quad\leqslant ||\Lambda_1^n - G_1^n - nG_1^{n-1}(\Lambda_1 - G_1)||\,||W_1|| + n|G_1^n(E - G)(\Lambda_1 - G_1)|\,||W_1||\\
&\qquad + n|G_1^n(\Lambda_1 - G_1 - A_2(E_1 - E)^3)|\,||W_1|| + ||G_1^n(W_1 - E - A_1(E_1 - E)\\
&\qquad + nA_2(E_1 - E)^3(W_1 - E))|| + ||\Lambda_2||^n||W_2||\\
&\quad\leqslant \sum_{j=2}^{n}(j-1)||\Lambda_1^{n-j}G_1^{j-2}(\Lambda_1 - G_1)^2|| + Cn\bar{q}^2(p+\bar{q})^2||G_1^{n-1}(E_1 - E)^4||\\
&\qquad + nC\bar{q}(p+\bar{q})^3||G_1^n(E_1 - E)^4|| + C|\bar{q} - p|(\bar{q} + p)||G_1^n(E_1 - E)^2||\\
&\qquad + CA_2n|\bar{q} - p|\,||G_1^n(E_1 - E)^4||.
\end{aligned} \tag{3.35}$$

The rest of the proof is a systematic application of Lemma 3.1 with $v = nv_1$, $v = (n-1)v_1$ or $v = (n-2)v_1$. □

4. Signed compound Poisson approximations

We consider the scheme of Section 3. Evidently, the estimates in Theorems 3.1–3.3 are trivial if $n\bar{q} = \lambda + o(1)$, $p_0 = o(1)$, and $p \equiv \text{Const}$. This is easy to explain. These conditions are sufficient for the existence of the limiting compound Poisson distribution with the compounding geometric distribution. Moreover, under these conditions F_n is close to its limiting distribution, but not to the normal or Poisson law. Consequently, SP substitutes for the normal and Poisson approximations do not fit. In this section, we consider an SCP approximation which is a universal replacement for the compound Poisson and normal approximations. Note that the Poisson distribution is only a partial case of the compound Poisson distribution.

The compound Poisson limit for sums of Markov dependent Bernoulli variables under slightly varying definitions of S_n was obtained by many authors, see, for example, Koopman (1950), Dobrushin (1953), Isham (1980), Wang (1981), Gani (1982), Brown (1983), Serfozo (1986), Wang and Bühler (1991), and references therein. Further developments in the theory of compound Poisson approximations can be found in Barbour et al. (1992), Roos (1994), and Geske et al. (1995).

We will present one of the estimates of the accuracy of compound Poisson approximation obtained by Wang (1992). The notation is that of Section 3. Let H_r be the geometric distribution,

$$\widehat{H}_r(t) = re^{it}(1 - (1-r)e^{it})^{-1}, \quad 0 \leqslant r \leqslant 1,$$

$S_{n1} = \xi_0 + \xi_1 + \cdots + \xi_{n-1}$, $\lambda > 0$ and let F_{n1} be the distribution of S_{n1}.

Theorem 4.1. (Wang, 1992). *Let* $0 < \bar{q}, q, p_0 \leqslant 1$. *Then*

$$\begin{aligned}
||F_{n1} - \exp\{\lambda(H_r - E)\}|| &\leqslant \max(p_0, \bar{q})(1 + 3n\bar{q}/q) + 2|(n-1)\bar{q} - \lambda|\\
&\quad + n\bar{q}|(1-q)^{K(r,q)} - (1-r)^{K(r,q)}|.
\end{aligned} \tag{4.1}$$

Here $K(r,q)$ *is the point where the expression* $1-((1-r)^{k-1}r)/(p^{k-1}q)$ *changes the sign.*

Estimate (4.1) is flexible because of possible different choices of λ and r. However, the rate of approximation in Eq. (4.1) is not better than $\bar{q}+n\bar{q}^2$, i.e., the estimate is sharp for small $\bar{q}$ only. Considering our setting we shall improve the rate of approximation for larger values of parameters. Here we consider only one of various possible compound Poisson laws.

Set

$$0\leqslant p\leqslant C_8<1, \tag{4.2}$$

$$H=qE_1\sum_{j=0}^{\infty}p^jE_j, \qquad (\widehat{H}(t)=q\mathrm{e}^{\mathrm{i}t}/(1-p\mathrm{e}^{\mathrm{i}t})), \tag{4.3}$$

where the distribution F_n was defined in Section 3.

Theorem 4.2. *Let condition* (4.2) *be satisfied. Then*

$$\left\|F_n-\exp\left\{\frac{n\bar{q}q}{q+\bar{q}}(H-E)\right\}\right\| \leqslant C_9\max(p_0,\bar{q})\min(1,(n\bar{q})^{-1/2})+C_{10}\min(\bar{q},n\bar{q}^2)+C_{11}\mathrm{e}^{-C_{12}n}, \tag{4.4}$$

$$\left\|F_n-\exp\left\{\frac{n\bar{q}q}{q+\bar{q}}(H-E)\right\}p_0(E+u_1(H-E))\right\| \leqslant C_{13}\bar{q}(p+\bar{q})\min(1,(n\bar{q})^{-1/2})+C_{10}\min(\bar{q},n\bar{q}^2)+C_{11}\mathrm{e}^{-C_{12}n}, \tag{4.5}$$

where $u_1=q^2(p-\bar{q})/(q+\bar{q})^2$.

Remark 4.1. Clearly, estimate (4.4) is decreasing if $\bar{q}=\mathrm{o}(1)$ and the condition $p_0=\mathrm{o}(1)$ or $n\bar{q}\to\infty$ is satisfied. By adding meanwhile one member of asymptotics we obtain that, in Eq. (4.5), this new approximation is decreasing whenever $\bar{q}=\mathrm{o}(1)$ (independently of the behaviour of p_0). Obviously, such assumptions are weaker than required for the smallness of the estimate in Eq. (4.1).

Remark 4.2. If $\bar{q}=\mathrm{o}(n^{-1/2})$, then Eq. (4.5) is sharper than the Berry–Esseen estimate.

Now we shall consider SCP approximations. Set

$$a_1=\frac{q\bar{q}}{q+\bar{q}}, \qquad a_2=-\frac{q\bar{q}^2}{(q+\bar{q})^2}\left(p+\frac{q}{q+\bar{q}}\right),$$

$$a_3=\frac{q\bar{q}^2}{(q+\bar{q})^3}\left\{p^2\bar{q}+\frac{qp(2\bar{q}-q)}{q+\bar{q}}+\frac{2\bar{q}q^2}{(q+\bar{q})^2}\right\}, \tag{4.7}$$

$$u_2 = \frac{q\bar{q}(\bar{q}-p)}{(q+\bar{q})^2}, \qquad u_3 = a_3 + \frac{q^2\bar{q}^3}{3(q+\bar{q})^3}\left(1+2p+\frac{3q}{q+\bar{q}}\right), \tag{4.8}$$

$$G_3 = \exp\{a_1(H-E)+(a_2-a_1^2/2)(H-E)^2\}. \tag{4.9}$$

Unlike the situation in Section 3, we do not assume any condition on p, except (4.2), i.e., we are going to get the estimates for almost every p, not only for a small one. However, we assume that $\bar{q}$ is quite small. Set

$$\bar{q}/(q+\bar{q}) \leqslant (1-C_8)/30. \tag{4.10}$$

Of course, for $\bar{q} = o(1)$, condition (4.10) is satisfied when $n \to \infty$. However, SCP approximations considered below provide small estimates even for $\bar{q} \sim$ Const.

Theorem 4.3. *Let assumptions* (4.2) *and* (4.10) *be satisfied. Then*

$$\|F_n - G_3^n\| \leqslant C_{14}(p+\bar{q})\{\min(\sqrt{\bar{q}/n}, n\bar{q}^2) + \max(p_0,\bar{q})\min(1,(n\bar{q})^{-1/2}) + e^{-C_{12}n}\}, \tag{4.11}$$

$$\|F_n - G_3^n p_0(E+u_1(H-E))\| \leqslant C_{15}\{\min(\sqrt{\bar{q}/n}, n\bar{q}^2) + \bar{q}\min(1,(n\bar{q})^{-1/2}) + e^{-C_{16}n}\} \leqslant C_{17}n^{-1/2}, \tag{4.12}$$

$$\|F_n - G_3^n\{p_0(E+u_1(H-E)) + (1-p_0)(E+u_2(H-E)) + nu_3(H-E)^3\}\| \leqslant C_{18}n^{-1}. \tag{4.13}$$

Remark 4.3. Various approaches can be used with respect to the asymptotics in Eq. (4.12). We choose u_1, u_2, u_3, so that Eq. (4.12) be always at least of order $n^{-1/2}$ and Eq. (4.13) be always at least of order n^{-1}.

Proof of Theorems 4.2 and 4.3. The proof is quite similar to that of Theorem 3.1. We have

$$\widehat{B}(t) = \frac{1}{1+\bar{q}-pe^{it}} = \frac{1}{q+\bar{q}} + \frac{p(e^{it}-1)}{q+\bar{q}}\widehat{B}(t), \tag{4.14}$$

$$\widehat{B}(t) = \frac{q}{q+\bar{q}}\frac{1}{1-pe^{it}} - \frac{p\bar{q}}{q+\bar{q}}(\widehat{H}(t)-1)\widehat{B}(t). \tag{4.15}$$

Evidently, $|e^{it}-1| \leqslant 2|\widehat{H}(t)-1|$. Expanding $\widehat{\Lambda}_1(t)$ as in Eq. (3.16) and using recursively Eqs. (4.14) and (4.15) we obtain

$$\widehat{\Lambda}_1(t) = 1 + \sum_{j=1}^{3} a_j(\widehat{H}(t)-1)^j + \theta C\bar{q}^3(p+\bar{q})|\widehat{H}(t)-1|^4, \tag{4.16}$$

$$\widehat{\Lambda}_1'(t) = \widehat{H}'(t)\sum_{j=1}^{3} ja_j(\widehat{H}(t)-1)^{j-1} + \theta C\bar{q}^3(p+\bar{q})|\widehat{H}(t)-1|^3. \tag{4.17}$$

For all t, we have

$$|\widehat{H}(t)-1|^2 \leqslant 2|Re\,\widehat{H}(t)-1|. \tag{4.18}$$

Let $|t|\leqslant\pi$. Then by Eqs. (4.2), (4.10), (4.15), and (4.16) we get $|\widehat{Y}(t)|\leqslant 4/15$ and, consequently,

$$\begin{aligned}|\widehat{\Lambda}_1(t)| &\leqslant |1+\widehat{Y}(t)/(4\widehat{B}(t))| + 15|\widehat{Y}(t)|^2/(16\cdot 11|\widehat{B}(t)|)\\ &\leqslant |1+a_1(\widehat{H}(t)-1)| + 4\bar{q}^2(q+\bar{q}^{-2}|Re\,\widehat{H}(t)-1|\\ &\quad +30\bar{q}^2|Re\,\widehat{H}(t)-1|(11(q+\bar{q}))^{-1}(1-\bar{q}/q)^{-2}\\ &\leqslant 1+\frac{\bar{q}}{q+\bar{q}}(Re\,\widehat{H}(t)-1)\left(1-C_8-\frac{5\bar{q}}{q+\bar{q}}-\frac{15^2\cdot 60\bar{q}}{11\cdot 14^2(q+\bar{q})}\right)\\ &\leqslant 1+C\bar{q}(Re\,\widehat{H}(t)-1)\leqslant \exp\{C\bar{q}(Re\,\widehat{H}(t)-1)\}.\end{aligned}\tag{4.19}$$

Similarly,

$$|\widehat{G}_3(t)|\leqslant \exp\{C\bar{q}(Re\,\widehat{H}(t)-1)\}.\tag{4.20}$$

Note also that Eqs. (4.2) and (4.11) are sufficient for obtaining

$$\begin{aligned}\|\Lambda_2\| &\leqslant \|p+\bar{q}\| + \frac{\|Y\|}{2}\,\frac{1}{1-4/15}\leqslant |p+\bar{q}| + \frac{30\bar{q}}{11(q+\bar{q})}\\ &\leqslant C_8+(1-C_8)/15+(1-C_8)/11\leqslant (1+4C_8)/5<1.\end{aligned}$$

Therefore

$$\|\Lambda_2\|^n\leqslant e^{-Cn}.\tag{4.21}$$

Quite similarly we establish

$$\Lambda_1-G_3-u_3(H-E)^3=V_1(H-E)^3\bar{q}^2(p+\bar{q}),\tag{4.22}$$

$$W_1-p_0(E+u_1(H-E))=V_2\bar{q}(p+\bar{q})(H-E),\qquad W_2=V_3(p+\bar{q}).\tag{4.23}$$

Here V_i are measures satisfying $\|V_i\|\leqslant C$, $i=1,2,3$. For brevity, we omit t in the equations below. By standard calculations we obtain

$$|\widehat{\Lambda}_1-\widehat{G}_3|\leqslant C\bar{q}^2(p+q)|\widehat{H}-1|^3,\qquad |\widehat{\Lambda}_1'-\widehat{G}_3'|\leqslant C\bar{q}^2(p+q)|\widehat{H}-1|^2,$$

$$|\widehat{\Lambda}_1-\exp\{a_1(\widehat{H}-1)\}|\leqslant C\bar{q}^2|\widehat{H}-1|^2,$$

$$|(\widehat{\Lambda}_1-\exp\{a_1(\widehat{H}-1)\})'|\leqslant C\bar{q}^2|\widehat{H}-1|,$$

$$|(\widehat{\Lambda}\exp\{-\mathrm{i}tv_1)\})'|\leqslant C\bar{q}|\widehat{H}-1|,\qquad |(\widehat{G}_3\exp\{-\mathrm{i}tv_1)\})'|\leqslant C\bar{q}|\widehat{H}-1|,$$

$$|(\exp\{a_1(\widehat{H}-1)-\mathrm{i}tv_1\})'|\leqslant C\bar{q}|\widehat{H}-1|,$$

$$\int_{-\pi}^{\pi}\exp\{Cn\bar{q}(Re\,\widehat{H}-1)\}\,\mathrm{d}t\leqslant C(n\bar{q})^{-1/2}.$$

Note that by properties of the variation norm

$$\begin{aligned}\|(H-E)\exp\{na_1(H_1-E)\}\| &\leqslant \|(E_1-E)\exp\{na_1(E_1-E)\}\|\\ &\leqslant C\min(1,(n\bar{q})^{-1/2}),\end{aligned}\tag{4.24}$$

where the last estimate can be obtained by applying Lemma 3.1. Similarly,

$$\begin{aligned}\| G_3^n(H-E)\| &\leqslant \|\exp\{a_1(E_1-E)+(a_2-a_1^2/2)(E_1-E)^2\}(E_1-E)\| \\ &\leqslant C\min(1,(n\bar{q})^{-1/2}). \end{aligned} \tag{4.25}$$

We have

$$\|F_n - G_3^n\| \leqslant \|\Lambda_1^n - G_3^n\|\ \|W_1\| + \|G_3^n(W_1-E)\| + \|\Lambda_2\|^n\ \|W_2\|,$$

$$\begin{aligned}&\|F_n - G_3^n p_0(E+u_1(H-E))\| \\ &\quad\leqslant \|F_n - G_3^n\|\ \|W_1\| + \|G_3^n(W_1 - p_0(E+u_1(H-E)))\| + \|\Lambda_2\|^n \|W_2\|.\end{aligned}$$

Analogous estimates hold for $\exp\{a_1(H-E)\}$. Note also that the left-hand side of Eq. (4.13) is less than or equal to

$$\begin{aligned}&\|\Lambda_1^n - G_3^n - nG_3^{n-1}(\Lambda_1 - G_3)\|\ \|W_1\| + n\|G_3^{n-1}(E-G_3)(\Lambda_1-G_3)\|\ \|W_1\| \\ &\quad + n\|G_3^n(\Lambda_1 - G_3 - u_3(H-E)^3)\|\ \|W_1\| + \|\Lambda_2\|^n\|W_2\| \\ &\quad + \|G_3^n(W_1 - p_0(E+u_1(H-E)) - (1-p_0)(E+u_2(H-E)) \\ &\quad + nu_3(H-E)^3(W_1-E))\|.\end{aligned}$$

Further the proofs of Theorems 4.2 and 4.3 are very similar to those of Theorems 3.1 and 3.3. Without loss of generality, we can assume that, in Theorem 4.2, condition (4.10) is satisfied and then repeatedly use the estimates obtained above and Lemma 3.1. □

5. Local estimates

Besides integral estimates, it is possible to obtain local ones. Note that such estimates are very natural, because we consider measures all concentrated on the integers. There is no need for discretization of the approximating measure (in contrast, e.g., to the situation in local theorems for the normal distribution). Evidently, local estimates can be obtained for the approximations of Sections 2 and 4.

Theorem 5.1. *Let assumptions* (2.6)–(2.8) *be satisfied. Then*

$$\sup_m |F_{n0}\{m\} - G_0^n\{m\}| = \mathrm{O}(n^{-1}), \tag{5.1}$$

$$\sup_m |F_{n0}\{m\} - G_0^n(E+G_{01})\{m\}| = \mathrm{o}(n^{-1}). \tag{5.2}$$

Proof. The inversion formula states that, for any finite measure Q is concentrated on the integers,

$$Q\{m\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}tm}\widehat{Q}(t)\,\mathrm{d}t, \tag{5.3}$$

whence

$$|F_{n0}\{m\} - G_0^n(E+G_{01})\{m\}| = \frac{1}{2\pi}\left|\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}tm}(\widehat{F}_{n0}(t) - \widehat{G}_0^n(t)(1+\widehat{G}_{01}(t))\,\mathrm{d}t\right|$$
$$\leqslant \frac{1}{\sqrt{n}}\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} |\widehat{\Delta}_0(t)|\,\mathrm{d}t.$$

The rest of the proof now coincides with that of Theorem 2.2. □

Theorem 5.2. *Let assumptions* (4.2) *and* (4.10) *be satisfied. Then*

$$\sup_m |F_n\{m\} - G_3^n p_0(E + u_1(H-E))\{m\}| \leqslant C_{19} n^{-1}. \tag{5.4}$$

Proof. The left-hand side of Eq. (5.4) is less than or equal to

$$\sup_m |\Lambda_1^n\{m\} - G_3^n\{m\}|\,\|W_1\|$$
$$+ \sup_m |G_3^n(W_1 - p_0(E+u_1(H-E)))\{m\}| + \|\Lambda_2\|^n\,\|W_2\|.$$

By Eq. (5.3) and the results obtained in the proof of Theorem 4.3 we easily deduce (5.4). □

The main part of this section is devoted to the local estimates depending on m. We further assume that m is an integer and $0 \leqslant m < n$. Consider the Markov binomial scheme of Sections 3 and 4. We shall show that, for any combination of p and $\bar{q}$, there exist SP approximations close to $P(S_n = m)$. For approximation we would apply m and $n-m$ convolutions of the approximating SP measures not depending on m in any other way. For brevity and convenience we assume that $p_0 = 1$ and denote $F_{n+1}\{m\} = P(S_{n+1} = m)$. Using the explicit expression in binomial coefficients of $P(S_{n+1} = m+1)$ from Dobrushin (1953) the following inversion formula can be obtained:

$$F_{n+1}\{m\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} (q + \bar{q}\mathrm{e}^{\mathrm{i}t})(q\mathrm{e}^{-\mathrm{i}t} + p)^m(\bar{p} + \bar{q}\mathrm{e}^{\mathrm{i}t})^{n-m}\,\mathrm{d}t. \tag{5.5}$$

(We are grateful to A. Bikelis who brought this fact to our attention). Set

$$G_4 = \exp\left\{\left(p - \frac{p^2}{2}\right)(E_1 - E) - \frac{p^2}{2}(E_{-1} - E)\right\},$$

$$G_5 = \exp\left\{\left(\bar{q} - \frac{\bar{q}^2}{2}\right)(E_1 - E) - \frac{\bar{q}^2}{2}(E_{-1} - E)\right\},$$

$$G_6 = \exp\left\{\left(\bar{p} - \frac{\bar{p}^2}{2}\right)(E_{-1} - E) - \frac{\bar{p}^2}{2}(E_1 - E)\right\},$$

$$G_7 = \exp\left\{\left(\bar{q} - \frac{\bar{q}^2}{2}\right)(E_{-1} - E) - \frac{\bar{q}^2}{2}(E_1 - E)\right\}.$$

Theorem 5.3. *Let* $p_0=1$. *Then for all* $n\geqslant 1$, $1<m<n$

(a) *if* $p\leqslant C_{20}<1$, $\bar q\leqslant C_{20}<1$ *then*

$$|F_{n+1}\{m\}-(q+\bar q)G_4^m G_5^{n-m}\{m\}|\leqslant \frac{C_{21}\max(p,\bar q)}{\max(1,mp+(n-m)\bar q)}\leqslant \frac{C_{22}}{n}\frac{\max(p,\bar q)}{\min(p,\bar q)},$$

(b) *if* $p\leqslant C_{20}<1$, $\bar p\leqslant C_{20}<1$ *then*

$$|F_{n+1}\{m\}-(qE_{n-m}G_4^mG_6^{n-m}+\bar qE_{n-m+1}G_4^mG_6^{n-m})\{m\}|$$
$$\leqslant \frac{C_{23}\max(p,\bar p)}{\max(1,mp+(n-m)\bar p)}\leqslant\frac{C_{24}}{n}\frac{\max(p,\bar p)}{\min(p,\bar p)},$$

$$|F_{n+1}\{m\}-(q+\bar q)E_{n-m}G_4^mG_6^{n-m}\{m\}|\leqslant\frac{C_{25}}{\max(1,mp+(n-m)\bar p)},$$

(c) *if* $q\leqslant C_{20}<1$, $\bar q\leqslant C_{20}<1$ *then*

$$|F_{n+1}\{m\}-(q+\bar q)E_mG_7^mG_5^{n-m}\{m\}|$$
$$\leqslant\frac{C_{26}\max(q,\bar q)}{\max(1,mq+(n-m)\bar q)}\leqslant\frac{C_{27}}{n}\frac{\max(q,\bar q)}{\min(q,\bar q)},$$

(d) *if* $q\leqslant C_{20}<1$, $\bar p\leqslant C_{20}<1$ *then*

$$|F_{n+1}\{m\}-(q+\bar q)E_{n+1}G_7^mG_6^{n-m}\{m\}|$$
$$\leqslant\frac{C_{28}\max(q,\bar p)}{\max(1,mq+(n-m)\bar p)}\leqslant\frac{C_{29}}{n}\frac{\max(q,\bar p)}{\min(q,\bar p)}.$$

Remark 5.1. Note that the local estimates presented above are sharp whenever the parameters ($p,\bar q$ etc.) are of the same order.

Remark 5.2. Note that all C in Theorem 5.3 do not depend on m.

Remark 5.3. Obviously, using the fact that $Q\{m+a\}=(E_{-a}Q)\{m\}$ we can reformulate all results replacing $E_{n-m}G_4^mG_6^{n-m}\{m\}$ by $G_4^mG_6^{n-m}\{2m-n\}$ etc.

Proof. We have

$$|(q+p\mathrm{e}^{\mathrm{i}t})-\widehat G_4(t)|\leqslant Cp^2|\sin(t/2)|^3,\qquad |(\bar p+\bar q\mathrm{e}^{\mathrm{i}t})-\widehat G_5(t)|\leqslant C\bar q^2|\sin(t/2)|^3,$$

$$|q+p\mathrm{e}^{\mathrm{i}t}|,\ |\widehat G_4(t)|\leqslant\exp\{-Cp\sin^2(t/2)\},$$

$$|\bar p+\bar q\mathrm{e}^{\mathrm{i}t}|,\ |\widehat G_5(t)|\leqslant\exp\{-C\bar q\sin^2(t/2)\}.$$

Then by Eqs. (5.3) and (5.5)

$$|F_{n+1}\{m\}-(q+\bar q)G_4^mG_5^{n-m}\{m\}|$$
$$\leqslant\frac{1}{2\pi}\int_{-\pi}^{\pi}\Big\{|q+\bar q\mathrm{e}^{\mathrm{i}t}|\,|\mathrm{e}^{-\mathrm{i}tm}|\,|((q+p\mathrm{e}^{\mathrm{i}t})^m-\widehat G_4^m(t))(\bar p+\bar q\mathrm{e}^{\mathrm{i}t})^{n-m}$$
$$+\widehat G_4^m(t)((\bar p+\bar q\mathrm{e}^{\mathrm{i}t})^{n-m}-\widehat G_5^{n-m}(t))|+\bar q|\widehat G_4^m(t)\widehat G_5^{n-m}(t)|\,|\mathrm{e}^{\mathrm{i}t}-1|\Big\}\,\mathrm{d}t$$

$$\leqslant C\int_{-\pi}^{\pi} \exp\{-C(mp+(n-m)\bar{q})\sin^2(t/2)\}((mp^2+(n-m)\bar{q}^2)|\sin^3(t/2)|$$
$$+\bar{q}|\sin(t/2)|)\,\mathrm{d}t$$
$$\leqslant C\max(p,\bar{q})(\max(1,mp+(n-m)\bar{q}))^{-1/2}$$
$$\int_{-\pi}^{\pi} \exp\{-C(mp+(n-m)\bar{q})\sin^2(t/2)\}\,\mathrm{d}t$$
$$\leqslant C\max(p,\bar{q})(\max(1,mp+(n-m)\bar{q}))^{-1}.$$

All other estimates are obtained by the same way. □

6. Concluding remarks

In this paper we considered theoretical aspects of the new SP approximations of discrete distributions. It seems that not only Poisson approximations should be modified due to the signed Poisson approach, but, in many cases, the classical normal approximations also should be revised.

What can be said about the computational aspects of SCP measures? Approximations G_1 and G_2 have simpler structures than G_3 and, therefore, are more convenient for practical applications. On the other hand, G_3 holds even when G_1 and G_2 fail. However, as it was shown by Hipp (1986), even G_3 can be applied in practice (the recursive formulae and numeric examples are given). There are other actuarial papers, where computational aspects of SP and SCP approximations are treated. For example, from the practical point of view SCP measures are discussed in Kuon et al. (1987). Note also that G_1 and G_2 have quite simple structures and can be expressed in Bessel functions and Hermite polynomials, respectively. Consequently, for their computation, in addition to recursions (see Kruopis 1986b, Borovkov and Pfeifer (1996)), the well-developed theory of special functions also can be applied.

One SP approximation is as good as the set of all limiting laws for the binomial distribution, see Eq. (1.3). In Section 5, we saw that, to some extent, the same can be said about the local SP approximations for the Markov binomial distribution. However, it is unclear, whether it is possible to construct an SCP approximation comparable with all limiting laws when the uniform or variation distances are considered.

The approximation G_3 is only one of the possibilities. For example, with the same rate of accuracy, G_3 can be replaced by the SCP approximation having the Fourier–Stieltjes transform

$$\exp\left\{\frac{\bar{q}(\mathrm{e}^{\mathrm{i}t}-1)}{1+\bar{q}-p\mathrm{e}^{\mathrm{i}t}} - \frac{\bar{q}^2(2+q+\bar{q})}{q+\bar{q}}\frac{(\mathrm{e}^{\mathrm{i}t}-1)^2}{(1+\bar{q}-p\mathrm{e}^{\mathrm{i}t})}\right\}.$$

As it follows from Theorems 3.3 and 4.3, all asymptotics should be constructed not only for Λ_1^n, but for W_1 as well.

In general, we used an operator technique which is very similar to that of Deheuvels and Pfeifer (1986). On the other hand, in the case we needed more precise estimates, we applied Lemma 3.1, i.e., we used the characteristic function method. However, the constants obtained were not reasonably small, therefore we concentrated our efforts only on the rates of approximations. A more precise estimation of constants probably would require methods less standard than those used in this paper. We only note that, for small values of $\bar{q}$ and p, the choice of parameters may be slightly different from ours and similar to that proposed by Borovkov and Pfeifer (1996). Such a choice cannot improve the rate of accuracy, but may result in smaller values of absolute constants.

We must also note that *asymptotical* constants are not large. For example, in Theorem 3.1, let $p_0 = 0$, $p = \text{Const} \leqslant 1/20$, $\bar{q} \to 0$, and $n\bar{q} \to \infty$ as $n \to \infty$. Then $\| F_n - G_1^n \| = \mathrm{O}((n\bar{q})^{-1/2})$. Applying quite standard calculations (just like in Prokhorov (1953)) we get

$$\lim_{n\to\infty} \sqrt{n\bar{q}} \| F_n - G_1^n \| = 0.190\ldots \frac{p^2}{(1+p)^{3/2}} \leqslant 0.000442\ldots .$$

In Section 4, we mentioned that the compound Poisson limit for the Markov binomial distribution was obtained by numerous authors. Different methods were used for this purpose. Here we give one more method based on a *local approach*. Consider the Markov binomial distribution from Section 3 with $p_0 = 1$. The last assumption means that, the compound Poisson limit with the compounding geometric distribution, cannot be a limiting law, see Eqs. (4.1) and (4.4). Which limit can we expect in this situation? The answer can be obtained from Eq. (4.5) (or, e.g. from Dobrushin's (1953) paper). However, we think that the following local approach is of its own interest. Let H_r and H be the same as in Section 4, i.e., $\widehat{H}(t) = q\mathrm{e}^{\mathrm{i}t}/(1 - p\mathrm{e}^{\mathrm{i}t})$ and $\widehat{H}_r(t) = r\mathrm{e}^{\mathrm{i}t}/(1-(1-r)\mathrm{e}^{\mathrm{i}t})$, $0 < r \leqslant 1$. It is easy to verify that, for all integers $k \geqslant 0$, $j \geqslant 0$,

$$\frac{q}{2\pi}\int_{-\pi}^{\pi} (q + p\mathrm{e}^{\mathrm{i}t})^j \mathrm{e}^{\mathrm{i}t(k-j)}\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\pi}^{\pi} \widehat{H}^{k+1}(t)\mathrm{e}^{-\mathrm{i}t(j+1)}\,\mathrm{d}t. \tag{6.1}$$

Let φ_k, $k = 1, 2$, be defined by

$$\varphi_k(x) = \sum_{j=0}^{\infty} \beta_{kj} x^j, \qquad \sum_{j=0}^{\infty} |\beta_{kj}| < \infty. \tag{6.2}$$

Then by Eq. (6.1) we get the relation

$$\frac{q}{2\pi}\int_{-\pi}^{\pi} \varphi_1(p + q\mathrm{e}^{-\mathrm{i}t})\varphi_2(\mathrm{e}^{\mathrm{i}t})\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}t}\widehat{H}(t)\varphi_1(\mathrm{e}^{-\mathrm{i}t})\varphi_2(\widehat{H}(t))\,\mathrm{d}t. \tag{6.3}$$

Applying Eq. (6.3) to Eq. (5.5) we obtain the following inversion formula for all $0 \leqslant m < n$:

$$F_n\{m\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \widehat{H}(t)\mathrm{e}^{-\mathrm{i}t}(1 + \bar{q}\widehat{H}(t)/q)(\bar{p} + \bar{q}\widehat{H}(t))^{n-m-1}\mathrm{e}^{-\mathrm{i}tm}\,\mathrm{d}t. \tag{6.4}$$

From Eq. (6.4) it is evident that, indeed, F_n can be close to some compound distribution. We formulate this fact as follows.

Proposition. *Let $n\bar{q} \to \lambda$ and $q \to r > 0$ as $n \to \infty$. Then*

$$\| F_n - E_{-1}H_r \exp\{\lambda(H_r - E)\} \| \to 0 \quad \textit{as } n \to \infty. \tag{6.5}$$

Proof. From the results of Section 3 it is easy to obtain that, under the conditions of proposition,

$$\mathbf{E}S_n = \frac{(n+1)\bar{q}}{q+\bar{q}} + \frac{q}{(q+\bar{q})^2}(1-(p-\bar{q})^{n+1}) - 1 = \mathrm{O}(1). \tag{6.6}$$

Therefore by Chebyshev's inequality we have, for $m > n^{1/2}$,

$$F_n\{(m,\infty)\} \leqslant \mathbf{E}|S_n|/m = \leqslant \mathbf{E}|S_n|n^{-1/2} = \mathrm{o}(1). \tag{6.7}$$

Analogous relation holds for $E_{-1}H_r \exp\{\lambda(H_r - E)\}$. Now let $m \leqslant n^{1/2}$. By the inversion formulas (6.4) and(5.3) we get (we skip the dependence on t in the Fourier transforms):

$$\begin{aligned}
&|F_n\{m\} - E_{-1}H_r \exp\{\lambda(H_r - E)\}\{m\}| \\
&\quad \leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} (|(\bar{p}+\bar{q}\widehat{H})^{n-m-1}\mathrm{e}^{-\mathrm{i}t(m+1)}|(|\bar{q}\widehat{H}^2/q| + |\widehat{H} - \widehat{H}_r|) \\
&\qquad + |\widehat{H}_r\mathrm{e}^{-\mathrm{i}t(m+1)}|(|(\bar{p}+\bar{q}\widehat{H})^{n-m-1} - \exp\{(n-m-1)\,\bar{q}\,(\widehat{H}-1)\}| \\
&\qquad + |\exp\{(n-m-1)\,\bar{q}\,(\widehat{H}-1)\}|\,|1 - \exp\{(m+1)\,\bar{q}\,(\widehat{H}-1)\}| \\
&\qquad + |\exp\{n\bar{q}(\widehat{H}-1)\} - \exp\{\lambda(\widehat{H}-1)\}| + |\exp\{\lambda(\widehat{H}-1)\} \\
&\qquad - \exp\{\lambda(\widehat{H}_r - 1)\}|))\,\mathrm{d}t.
\end{aligned} \tag{6.8}$$

It is easy to show that the right-hand side of Eq. (6.8) is o(1) as $n \to \infty$. Therefore by the Scheffe dominant convergence theorem we obtain (6.5). □

Acknowledgements

We are grateful to V. Kazakevičius for useful discussions. We thank the anonymous referees for many valuable remarks.

References

Barbour, A.D., Holst, L., Janson, S., 1992. Poisson Approximations. Oxford University Press, Oxford.

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

Borovkov, K.A., 1988. Refinement of Poisson approximation. Theory Probab. Appl. 33, 343–347.

Borovkov, K.A., Pfeifer, D., 1996. Pseudo-Poisson approximation for Markov chains. Stochastic Process. Appl. 61, 163–180.

Brown, T.C., 1983. Some Poisson approximation using compensators. Ann. Probab. 11, 726–744.

Čekanavičius, V., 1996. On multivariate Le Cam theorem and compound Poisson measures. Statist. Probab. Lett. 25, 145–151.

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

Čekanavičius, V., 1998. On signed normal-Poisson approximations. Probab. Theory Related Fields 111, 565–583.

Cuppens, R., 1975. Decomposition of Multivariate Probability. Academic Press, New York.

Deheuvels, P., Pfeifer, D., 1986. A semigroup approach to Poisson approximation. Ann. Probab. 14, 663–676.

Dobrushin, R.L., 1953. Limit theorems for a Markov chain of two states. Izv. Akad. Nauk. USSR Ser. mat. 17, 291–330 (in Russian).

Gani, J., 1982. On the probability generating function of the sum of Markov–Bernoulli random variables. J. Appl. Probab. 19 A, 321–326.

Geske, M.X., Godbole, A.P., Schafner, A.A., Skolnick, A.M., Wallstrom, G.L., 1995. Compound Poisson approximations for word patterns under Markovian hypotheses. J. Appl. Probab., 32.

Gnedenko, B.V., Kolmogorov, A.N., 1954. Limit Distributions for Sums of Independent Random Variables. Addison-Wesley, Reading, MA.

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

Isham, I., 1980. Dependent thinning of point proceses. J. Appl. Probab. 17, 987–995.

Koopman, B.O., 1950. A generalization of Poisson's distribution for Markoff chains. Proc. Natl. Acad. Sci. USA 36, 202–207.

Kornya, P., 1983. Distribution of aggregate claims in the Individual Risk Theory model. Soc. Actuaries: Trans. 35, 823–858.

Kruopis, J., 1986a. Precision of approximations of the generalized binomial distribution by convolutions of Poisson measures. Lithuanian Math. J. 26, 37–49.

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

Kuon, S., Reich, A., Reimers, L., 1987. Panjer vs. Kornya vs. De Pril: a comparison from a practical point of view. ASTIN Bull. 17 (2), 183–191.

Presman, E.L., 1983. Approximation of binomial distributions by infinitely divisible ones. Theory Probab. Appl. 28, 393–403.

Presman, E.L., 1985. The variation distance between the distribution of a sum of independent Bernoulli variables and the Poisson law. Theory Probab. Appl. 30, 417–422.

Prokhorov, Yu.V., 1953. Asymptotic behavior of the binomial distribution. Uspekhi Mat. Nauk 8 (3), 135–142 (in Russian).

Roos, M., 1994. Stein's method for compound Poisson approximation: the local approach. Ann. Appl. Probab. 4, 1177–1187.

Serfling, R.J., 1975. A general Poisson approximation theorem. Ann. Probab. 3, 726–731.

Serfozo, R.F., 1986. Compound Poisson approximation for sums of random variables. Ann. Probab. 14, 1391–1398: Correction Ann. Probab. 16 (1988) 429.

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

Sirazhdinov, S.H., Formanov, Sh.K., 1979. Limit theorems for sums of random vectors forming a Markov chain. Tashkent (in Russian).

Sirazhdinov, S.H., 1952. Refinement of limit theorems for homogeneous Markov chains. Dokl. Akad. Nauk SSSR 84 (6) (in Russian).

Wang, Y.H., 1981. On the limit of the Markov binomial distribution. J. Appl. Probab. 18, 937–942.

Wang, Y.H., 1992. Approximating kth order two-state Markov chains. J. Appl. Probab. 29, 861–868.

Wang, Y.H., Bühler, W.J., 1991. Renewal process proof for the limit of the Markov binomial distribution. Math. Scientist 16, 66–68.