UNIVERSITÄT HAMBURG

Compound binomial approximations

Vydas Čekanavičius and Bero Roos

Preprint–No. 2005–02 March 2005

FACHBEREICH MATHEMATIK

SCHWERPUNKT MATHEMATISCHE STATISTIK
UND STOCHASTISCHE PROZESSE

BUNDESSTR. 55, D – 20146 HAMBURG

COMPOUND BINOMIAL APPROXIMATIONS

VYDAS ČEKANAVIČIUS[1] AND BERO ROOS[2]

[1]*Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, Vilnius 03225, Lithuania, e-mail: vydas.cekanavicius@maf.vu.lt*

[2]*Corresponding author. Department of Mathematics, SPST, University of Hamburg, Bundesstr. 55, 20146 Hamburg, Germany, e-mail: roos@math.uni-hamburg.de*

Revised version, March 10, 2005

Abstract. We consider the approximation of the convolution product of not necessarily identical probability distributions q_jI+p_jF, $(j=1,\dots,n)$, where, for all j, $p_j=1-q_j\in[0,\,1]$, I is the Dirac measure at point zero, and F is a probability distribution on the real line. As an approximation, we use a compound binomial distribution, which is defined in a one-parametric way: the number of trials remains the same but the p_j are replaced with their mean or, more generally, with an arbitrary success probability p. We also consider approximations by finite signed measures derived from an expansion based on Krawtchouk polynomials. Bounds for the approximation error in different metrics are presented. If F is a symmetric distribution about zero or a suitably shifted distribution, the bounds have a better order than in the case of a general F. Asymptotic sharp bounds are given in the case, when F is symmetric and concentrated on two points.

Key words and phrases: Compound binomial distribution, Kolmogorov norm, Krawtchouk expansion, concentration norm, one-parametric approximation, sharp constants, shifted distributions, symmetric distributions, total variation norm.

1. Introduction

1.1 *The aim of the paper*

Binomial approximations are usually applied to the sums of independent or dependent indicators (see Ehm (1991), Barbour *et al.* (1992, Section 9.2), Soon (1996), Roos (2000), Čekanavičius and Vaitkus (2001), Choi and Xia (2002)). The present paper is devoted to a more general situation of a generalized Poisson binomial distribution, which is defined as the convolution product of not necessarily identical probability distributions $q_jI + p_jF$, $(j \in \{1, \dots, n\},\ n \in \mathbb{N} = \{1, 2, \dots\})$ on the set of real numbers $\mathbb{R}$. Here, for all j, $p_j = 1 - q_j \in [0,\, 1]$, I_u is the Dirac measure at point $u \in \mathbb{R}$, $I = I_0$, and F is assumed to be a probability distribution on $\mathbb{R}$. As approximations, we choose a compound binomial distribution, which is defined in a one-parametric way: the number n of trials remains the same but the p_j are replaced with their mean or, more generally, with an arbitrary success probability p. We also consider the approximation with a suitable finite signed measure, which can be derived from a related expansion based on Krawtchouk polynomials. Bounds for the distance in several metrics are given. It turns out that, if F is a symmetric distribution about zero or a suitably shifted distribution, i.e. the convolution of a distribution G on $\mathbb{R}$ with an adequate Dirac measure I_u at point $u \in \mathbb{R}$, the accuracy of approximation will be increased. In the compound Poisson approximation, similar investigations were made by Le Cam (1965) and Arak and Zaĭtsev (1988). In the case of a symmetric distribution F concentrated on two points, we present bounds containing asymptotic sharp constants.

Remarkably, the main body of research in compound approximations is restricted to compound Poisson laws or finite signed measures derived from related expansions. It seems that compound binomial approximation hardly attracted any attention. It is evident that, in contrast to approximations by (compound) Poisson laws, the ones by (compound) binomial distributions are exact when $p_j = p$ for all j. As can be seen below, the bounds for the distances given in this paper reflect this fact. It should be noted that the one-parametric binomial approximation is not the only one in this context. In Barbour *et al.* (1992, Section 9.2), it was proposed to use the two-parametric binomial approximation, matching two moments of the approximated law. The two-parametric approach can also be extended to the compound case and will be discussed in a separate paper (see Čekanavičius and Roos (2004)).

In this paper, a combination of different techniques is used. The main arguments are the Krawtchouk expansion from Roos (2000) and several norm estimates. It should be mentioned that the proofs of some of these norm estimates are rather complicated and require the deep Arak-Zaĭtsev method (cf. Čekanavičius (1995) and Arak and Zaĭtsev (1988)).

The structure of the paper is the following. In the next subsection, we proceed with some notation. In Section 2, we discuss important known facts. Section 3 is devoted to the main results, the proofs of which are given in Section 4. In an appendix, we collect some important facts on the Krawtchouk expansion.

1.2 *Some notation*

For our purposes, it is more convenient to formulate all results in terms of distributions or signed measures rather than in terms of random variables. Let $\mathcal{F}$ (resp. $\mathcal{S}$, resp. $\mathcal{M}$) denote the set of all probability distributions (resp. symmetric probability distributions about zero, resp. finite signed measures) on $\mathbb{R}$. All products and powers of finite signed measures in $\mathcal{M}$ are defined in the convolution sense; for $W \in \mathcal{M}$, set $W^0 = I$. The exponential of W is defined by the finite signed measure

$$\exp\{W\} = \sum_{m=0}^{\infty} \frac{W^m}{m!}.$$

Let $W = W^+ - W^-$ denote the Hahn-Jordan decomposition of W. The total variation norm $\|W\|$, the Kolmogorov norm $|W|$, and the Lévy concentration norm $|W|_h$ of $W \in \mathcal{M}$ are defined by

$$\begin{aligned}
\|W\| &= W^+(\mathbb{R}) + W^-(\mathbb{R}), \\
|W| &= \sup_{x\in\mathbb{R}} |W((-\infty, x])|, \\
|W|_h &= \sup_{x\in\mathbb{R}} |W([x,\, x+h])|, \qquad (h \in [0,\, \infty)),
\end{aligned}$$

respectively. Note that the total variation distance $\|F - G\|$ between $F, G \in \mathcal{F}$ is equal to $2\sup_A |F(A) - G(A)|$, where the supremum is over all Borel measurable sets $A \subseteq \mathbb{R}$. It should be mentioned that $|\cdot|_0$ is only a seminorm on $\mathcal{M}$, i.e. it may happen that, for non-zero $W \in \mathcal{M}$, $|W|_0 = 0$. But if we restrict ourselves to finite signed measures concentrated on the set of all integers $\mathbb{Z}$, then $|\cdot|_0$ is indeed a norm, the so-called local norm, that coincides with the ℓ_∞-norm of the counting density of the signed measure under consideration. We denote by C positive absolute constants, which may differ from line to line. Similarly, by $C(\cdot)$ we denote constants depending on the indicated argument only. By a condition of the type $f(x) \leq C < 1$ for a real-valued function $f(x)$ of some values x, we mean that a positive absolute constant $C < 1$ exists such that, for all x, $f(x) \leq C$. In other words, $f(x)$ is bounded away from 1 uniformly in x. For $x \in \mathbb{R}$, let $\lfloor x \rfloor$ be the largest integer not exceeding x. We always let $0^0 = 1$,

$$n \in \mathbb{N}, \qquad p_j \in [0,\, 1], \qquad q_j = 1 - p_j, \qquad (j \in \{1, \dots, n\}),$$

$$\mathbf{p} = (p_1, \dots, p_n), \qquad p_{\max} = \max_{1\leq j\leq n} p_j, \qquad p_{\min} = \min_{1\leq j\leq n} p_j, \qquad \delta = p_{\max} - p_{\min},$$

$$\overline{p} = \frac{1}{n}\sum_{j=1}^{n} p_j, \qquad \overline{q} = 1 - \overline{p}, \qquad \lambda = n\overline{p}, \qquad p \in [0,\, 1], \qquad q = 1 - p,$$

$$\gamma_k(p) = \sum_{j=1}^{n} (p - p_j)^k, \qquad \gamma_k = \gamma_k(\overline{p}), \qquad (k \in \mathbb{N}),$$

$$\eta(p) = 2\gamma_2(p) + (\gamma_1(p))^2, \qquad \theta(p) = \frac{\eta(p)}{2npq}, \qquad \theta = \theta(\overline{p}) = \frac{\gamma_2}{n\overline{p}\,\overline{q}},$$

$$\mathrm{GPB}(n,\, \mathbf{p},\, F) = \prod_{j=1}^{n}(q_j I + p_j F), \qquad \mathrm{Bi}(n,\, p,\, F) = (qI + pF)^n, \qquad (F \in \mathcal{F}).$$

Note that $\gamma_k(p)$, $\eta(p)$, and $\theta(p)$ not only depend on p but also on $\mathbf{p}$. For brevity, this dependence will not be explicitly indicated. The binomial distribution with parameter n and p is defined by $\mathrm{Bi}(n,\, p) = \mathrm{Bi}(n,\, p,\, I_1)$.

Using the above notation, the goal of the present paper can be summarized as follows: Give bounds for the accuracy of approximation of the generalized Poisson binomial distribution $\mathrm{GPB}(n,\, \mathbf{p},\, F)$, $(F \in \mathcal{F})$ by the compound binomial law $\mathrm{Bi}(n,\, p,\, F)$ and by related finite signed measures, which are defined in Subsection 2.2 below.

2. Known facts

2.1 *Ehm's result*

In what follows, we discuss some known results in the one-parametric binomial approximation to the Poisson binomial distribution $\mathrm{GPB}(n,\, \mathbf{p},\, I_1)$. By using Stein's method, Ehm (1991) proved that the total variation distance $d_1 = \|\mathrm{GPB}(n,\, \mathbf{p},\, I_1) - \mathrm{Bi}(n, \overline{p})\|$ between $\mathrm{GPB}(n,\, \mathbf{p},\, I_1)$ and the binomial distribution $\mathrm{Bi}(n,\, \overline{p})$ can be estimated in the following way:

$$\frac{1}{62}\min\{\theta,\, \gamma_2\} \le d_1 \le 2\min\{\theta,\, \gamma_2\}. \tag{2.1}$$

From (2.1), we see that d_1 and $\min\{\theta,\, \gamma_2\}$ have the same order. In estimating d_1, the quantities θ and γ_2 play a different rôle. First note that, as has been shown in Roos (2000, Remark on page 259), we have

$$\theta \le \delta \min\Big\{1,\, \frac{\delta}{4\overline{p}\,\overline{q}}\Big\}. \tag{2.2}$$

In particular, we have $\theta \le 1$. Since $\theta \le \delta \le \sum_{j=1}^{n} |\overline{p} - p_j| \le 2n\overline{p}\,\overline{q}$ (cf. Roos (2000, page 263)), we obtain

$$\frac{\theta^2}{2} \le \min\{\gamma_2, \theta\} \le \theta,$$

which implies that the distance d_1 is small if and only if θ is small, or, since $\theta = 1 - (n\overline{p}\,\overline{q})^{-1}\sum_{j=1}^{n} p_j q_j$, if and only if the quotient of the variances of the involved distributions is approximately equal to one (see also Ehm (1991, Corollary 2)). Therefore, looking at (2.1), the upper bound θ is much more important than the γ_2. To say it using the terminology by Barbour *et al.* (1992, page 5), the factor $\theta/\gamma_2 = (n\overline{p}\,\overline{q})^{-1}$ is a magic factor.

2.2 *Approximations using the Krawtchouk expansion*

In Roos (2000), the same binomial approximation problem as in Subsection 2.1 was investigated. In fact, by using generating functions, an expansion based on Krawtchouk polynomials was constructed. In what follows, we collect some basic facts about this expansion needed later on. Further informations can be found in the appendix below. In Theorem 1 of the above mentioned paper, it was shown that, for arbitrary $p \in [0, 1]$ and $F = I_1$, the identity

$$\mathrm{GPB}(n, \mathbf{p}, F) = \sum_{j=0}^{n} a_j(p)\,(F - I)^j\,(qI + pF)^{n-j} \tag{2.3}$$

holds, where the Krawtchouk coefficients $a_j(p)$ are given by

$$a_0(p) = 1, \qquad a_1(p) = -\gamma_1(p), \qquad a_2(p) = \frac{1}{2}((\gamma_1(p))^2 - \gamma_2(p)),$$
$$a_3(p) = -\frac{1}{6}(\gamma_1(p))^3 + \frac{1}{2}\gamma_1(p)\gamma_2(p) - \frac{1}{3}\gamma_3(p),$$

and, for $j \in \{1, \dots, n\}$,

$$a_j(p) = -\frac{1}{j}\sum_{k=0}^{j-1} a_k(p)\,\gamma_{j-k}(p). \tag{2.4}$$

Note that the coefficients $a_j(p)$ not only depend on p but also on $\mathbf{p}$. Alternatively, by (2.4), $a_j(p)$ can be considered as a function of $(\gamma_1(p), \dots, \gamma_j(p))$. It is evident that (2.3) also holds for a general distribution $F \in \mathcal{F}$. Taking into account (2.3), as an approximation of $\mathrm{GPB}(n, \mathbf{p}, F)$, it is useful to choose the finite signed measure

$$\mathrm{Bi}(n, p, F; s) = \sum_{j=0}^{s} a_j(p)\,(F - I)^j\,(qI + pF)^{n-j}, \qquad (F \in \mathcal{F}) \tag{2.5}$$

with $s \in \{0, \dots, n\}$ being fixed. Note that $\mathrm{Bi}(n, p, I_1; 0) = \mathrm{Bi}(n, p)$ and that, for $s = 1$ and $p = \overline{p}$, we have $\mathrm{Bi}(n, \overline{p}, I_1; 1) = \mathrm{Bi}(n, \overline{p})$. It should be mentioned that, in the remaining cases, $\mathrm{Bi}(n, p, F; s)$ also depends on $\mathbf{p}$. In Theorem 2 of Roos (2000), it was shown that

$$\|\mathrm{GPB}(n, \mathbf{p}, I_1) - \mathrm{Bi}(n, p, I_1; s)\| \le C_1(s)\min\{\theta(p), \eta(p)\}^{(s+1)/2}, \tag{2.6}$$
$$|\mathrm{GPB}(n, \mathbf{p}, I_1) - \mathrm{Bi}(n, p, I_1; s)|_0 \le C_2(s)\frac{(\theta(p))^{(s+1)/2}}{\sqrt{npq}}, \tag{2.7}$$

where, for (2.7), we have to assume that $\theta(p) \le C < 1$. For $p = \overline{p}$ and $s = 1$, (2.6) has the same order as Ehm's upper bound (see (2.1)). Note that the constants $C_1(s)$ and $C_2(s)$ can be given explicitly (see also Roos (2001a, Corollary 1)). In Roos (2000, Theorem 3), it was proved that, if $\gamma_2 > 0$, then, for $W = \mathrm{GPB}(n, \mathbf{p}, I_1) - \mathrm{Bi}(n, \overline{p})$,

$$\Big|\|W\| - \theta\sqrt{\frac{2}{\pi e}}\Big| \le C\,\theta\,v, \tag{2.8}$$
$$\Big||W|_0 - \frac{\theta}{2\sqrt{2\pi\, n\overline{p}\,\overline{q}}}\Big| \le C\,\frac{\theta}{\sqrt{n\overline{p}\,\overline{q}}}\,v, \tag{2.9}$$

with

$$v = \min\Big\{1, \frac{|\gamma_3|}{\gamma_2\sqrt{n\overline{p}\,\overline{q}}} + \frac{1}{n\overline{p}\,\overline{q}} + \theta\Big\},$$

where, for (2.9), we have to assume that $\theta < C < 1$. For example, from (2.8), it follows that $\|W\| \sim \sqrt{2/(\pi e)}\,\theta$ as $\theta \to 0$ and $n\overline{p}\,\overline{q} \to \infty$.

2.3 *The Le Cam-Michel trick and its generalizations*

As explained above, the purpose of the present paper is the investigation of the changes in the accuracy of approximation, when, in $\mathrm{GPB}(n,\, \mathbf{p},\, I_1)$ and $\mathrm{Bi}(n,\, p,\, I_1;\, s)$ from Subsection 2.2, I_1 is replaced with a more general distribution $F \in \mathcal{F}$. Thus we deal with compound distributions, which, however, retain structures similar to the ones presented in above. In fact, by the properties of the total variation norm,

$$(2.10)\quad \sup_{F\in\mathcal{F}} \|\mathrm{GPB}(n,\, \mathbf{p},\, F) - \mathrm{Bi}(n,\, p,\, F;\, s)\| = \|\mathrm{GPB}(n,\, \mathbf{p},\, I_1) - \mathrm{Bi}(n,\, p,\, I_1;\, s)\|.$$

Consequently, the supremum is achieved when $F = I_1$. In the compound Poisson approximation, a similar property has been observed by Le Cam (1965, page 187) and later rediscovered by Michel (1987, page 167). For the Kolmogorov norm and the concentration seminorm with $h = 0$, similar assertions hold. Moreover, generalizations with respect to arbitrary finite signed measures are possible. Indeed, if $W \in \mathcal{M}$ is concentrated on $\mathbb{Z}_+ = \{0, 1, 2, \ldots\}$ and if $F_1, F_2, F_3 \in \mathcal{F}$, where F_2 and F_3 are assumed to be concentrated on $[0,\, \infty)$ and on $(0,\, \infty)$, respectively, then, as is easily shown,

$$(2.11)\qquad \Big\|\sum_{m=0}^{\infty} W(\{m\})F_1^m\Big\| \le \|W\|,$$

$$(2.12)\qquad \Big|\sum_{m=0}^{\infty} W(\{m\})F_2^m\Big| \le |W| \sup_{x\in\mathbb{R}} \sum_{m=0}^{\infty}(F_2^m - F_2^{m+1})((-\infty,\, x]) \le |W|,$$

$$(2.13)\qquad \Big|\sum_{m=0}^{\infty} W(\{m\})F_3^m\Big|_0 \le |W|_0 \sup_{x\in\mathbb{R}} \sum_{m=0}^{\infty} F_3^m(\{x\}) \le |W|_0.$$

If, in (2.11)–(2.13), we set $W = \mathrm{GPB}(n,\, \mathbf{p},\, I_1) - \mathrm{Bi}(n,\, p,\, I_1;\, s)$, we arrive at (2.10) and the respective equalities for the Kolmogorov and local norms. In view of these inequalities, one can expect some improvement in approximation accuracy only under additional assumptions on F. Below, we will show that such improvements are possible, when F is either suitably shifted or symmetric. Note that, with a few exceptions only, we do not require the finiteness of moments of F.

2.4 *Compound Poisson approximations by Le Cam, Arak and Zaĭtsev*

In this paper, we consider shifted and symmetric distributions $F \in \mathcal{F}$. By a shifted distribution $F \in \mathcal{F}$, we mean $F = I_u G$, where $G \in \mathcal{F}$ and $u \in \mathbb{R}$. Then minimizing the distance with respect to u, one can expect some improvement of the accuracy of approximation. Shifted distributions play an important rôle in compound Poisson approximations, see, for example, Le Cam (1965) or Čekanavičius (2002). Indeed, Le Cam

(1965) proved that

$$\sup_{G\in\mathcal{F}}\inf_{u\in\mathbb{R}}\Big|(I_uG)^n-\exp\{n(I_uG-I)\}\Big|\le\frac{C}{n^{1/3}}. \tag{2.14}$$

Note that, as is easily shown, $\sup_{G\in\mathcal{F}}|G^n-\exp\{n(G-I)\}|$ is bounded away from zero, which stresses the advantage of shifted distributions in this context. Similar estimates hold for the symmetric distributions. More precisely, if in (2.14), we replace I_uG with an arbitrary symmetric distribution in $\mathcal{S}$ then the accuracy of compound Poisson approximation will be $Cn^{-1/2}$. If we replace I_uG with a symmetric distribution with non-negative characteristic function then the accuracy of compound Poisson approximation will be Cn^{-1}. In general, both results are of the right order, see Arak and Zaĭtsev (1988, Chapter 5).

It is noteworthy that, in a similar context, Barbour and Choi (2004) used Stein's method for the approximation of the sum of independent integer valued random variables by a translated Poisson distribution.

3. Main results

In what follows, let $\mathrm{Bi}(n,\,p,\,F;\,s)$, $(F\in\mathcal{F})$ be defined as in (2.5). Further, if not stated otherwise, we assume that

$$p\in[0,\,0.3]. \tag{3.1}$$

Theorem 3.1. *Let $s\in\{0,\dots,n\}$ and $h\in[0,\infty)$. Let us assume that*

$$\frac{4\theta(p)}{3(1-2p)^2}\le C<1. \tag{3.2}$$

Then the following inequalities hold. For $G\in\mathcal{F}$,

$$\inf_{u\in\mathbb{R}}|\mathrm{GPB}(n,\,\mathbf{p},\,I_uG)-\mathrm{Bi}(n,\,p,\,I_uG;\,s)|\le C(s)\,\frac{(\eta(p))^{(s+1)/2}}{(np)^{(s+1)/2+(s+1)/(2s+4)}}. \tag{3.3}$$

For $F\in\mathcal{S}$,

$$|\mathrm{GPB}(n,\,\mathbf{p},\,F)-\mathrm{Bi}(n,\,p,\,F;\,s)|\le C(s)\,\frac{(\eta(p))^{(s+1)/2}}{(np)^{s+1}}, \tag{3.4}$$

$$|\mathrm{GPB}(n,\,\mathbf{p},\,F)-\mathrm{Bi}(n,\,p,\,F;\,s)|_h\le C(s)\,\frac{(\eta(p))^{(s+1)/2}}{(np)^{s+1}}Q_h^{1/(2s+3)} \times(|\ln Q_h|+1)^{6(s+1)(s+2)/(2s+3)}, \tag{3.5}$$

where we write $Q_h:=Q_{h,\,np,\,F}:=|\exp\{32^{-1}np(F-I)\}|_h$.

Remark 1. (i) All approximations (3.3)–(3.5) are exact if $p_j=p$ for all j. In fact, in this case, the upper bounds vanish, since $\eta(p)=\gamma_2=0$. In particular, here, the conditions (3.1) and (3.2) are superfluous (see also Remark (ii)).

(ii) From (2.2), it follows that (3.2) is satisfied, when $p = \overline{p} \leq 0.3$ and $\delta \leq 0.3$, since, in this case, we have

$$\frac{4\theta}{3(1-2\overline{p})^2} \leq \frac{0.4}{(1-2\overline{p})^2} \min\Big\{1, \frac{0.3}{4\overline{p}\,\overline{q}}\Big\} \leq 0.9.$$

In particular, if $p_{\max} \leq 0.3$ and $p = \overline{p}$, then (3.2) is valid. In Theorem 3.1, the condition $p \in [0, 0.3]$ seems to be superfluous. It may suffice to assume that $\theta(p) \leq C$ for a suitable constant C. But we could not prove it.

(iii) The bounds (3.3)–(3.5) have a better order than $(\theta(p))^{(s+1)/2}$, which, in turn, appears as the interesting part in the bounds from (2.6) and (2.10). Generally, for the total variation distance, there seems to be no hope for upper bounds similar to (3.3) and (3.4). Good upper bounds for the concentration norm in the context of shifted distributions remain as an open question.

(iv) In contrast to (3.5), the estimates (3.3) and (3.4) are uniform in $G \in \mathcal{F}$ and $F \in \mathcal{S}$, respectively. However, due to the method of proof, we cannot say much about the constants $C(s)$. It should be mentioned, that, in order to obtain the explicit conditions (3.1) and (3.2), in the proofs we often deal with explicit constants. But, since our goal was to obtain a weak condition, the leading constants in the estimates turned out to be quite large.

(v) It is easily shown that (3.4) follows from (3.5). However, (3.5) leads to estimates of a better order than (3.4), if we use a Le Cam-type bound for the concentration function of compound Poisson distributions; for example, see Roos (2004, Proposition 3), where it was shown that, for $t \in (0, \infty)$, $h \in [0, \infty)$, and an arbitrary distribution $F \in \mathcal{F}$,

$$(3.6) \qquad |\exp\{t(F - I)\}|_h \leq \frac{1}{\sqrt{2e\,t \max\{F((-\infty, -h)),\, F((h, \infty))\}}}.$$

If, in Theorem 3.1, we set $p = \overline{p}$ and $s = 1$, we obtain the results with respect to the compound binomial approximation.

Corollary 3.1. *Let $h \in [0, \infty)$. Let us assume that*

$$(3.7) \qquad \overline{p} \in [0, 0.3] \quad \text{and} \quad \frac{4\theta}{3(1-2\overline{p})^2} \leq C < 1.$$

Then the following inequalities hold. For $G \in \mathcal{F}$,

$$\inf_{u \in \mathbb{R}} |\mathrm{GPB}(n, \mathbf{p}, I_u G) - \mathrm{Bi}(n, \overline{p}, I_u G)| \leq C\, \frac{\gamma_2}{\lambda^{4/3}}.$$

For $F \in \mathcal{S}$,

$$|\mathrm{GPB}(n, \mathbf{p}, F) - \mathrm{Bi}(n, \overline{p}, F)| \leq C\, \frac{\gamma_2}{\lambda^2},$$
$$|\mathrm{GPB}(n, \mathbf{p}, F) - \mathrm{Bi}(n, \overline{p}, F)|_h \leq C\, \frac{\gamma_2}{\lambda^2}\, Q_h^{1/5}(|\ln Q_h| + 1)^{36/5},$$

where Q_h is defined as in Theorem 3.1.

For symmetric distributions concentrated on $\mathbb{Z}\setminus\{0\}$, alternative estimates can be shown. In particular, in this case, it is possible to derive a bound for the total variation norm, which is comparable with (3.4) for the weaker Kolmogorov norm.

THEOREM 3.2. *Let the assumptions of Theorem 3.1 be valid. If $F \in \mathcal{S}$ is concentrated on the set $\mathbb{Z}\setminus\{0\}$, then*

$$\|\mathrm{GPB}(n,\,\mathbf{p},\,F) - \mathrm{Bi}(\,n,\,p,\,F;\,s)\| \leq C(s)\,\sqrt{\sigma}\,\frac{(\eta(p))^{(s+1)/2}}{(np)^{s+1}}, \tag{3.8}$$

$$|\mathrm{GPB}(n,\,\mathbf{p},\,F) - \mathrm{Bi}(n,\,p,\,F;\,s)|_h \leq C(s)\,\lfloor h+1\rfloor\,\frac{(\eta(p))^{(s+1)/2}}{(np)^{s+3/2}}, \tag{3.9}$$

where, for (3.8)*, we assume that F has finite variance σ^2.*

Remark 2. (i) The total variation bound (3.8) is slightly worse than (3.4); indeed, the variance σ^2 of $F \in \mathcal{S}$ concentrated on $\mathbb{Z}\setminus\{0\}$ cannot be smaller than one.

(ii) Under the assumptions of Theorem 3.2, an upper bound for the Kolmogorov norm can be shown by using Tsaregradskii's (1958) inequality. Unexpectedly, the resulting bound is of worse order than (3.8) and is therefore omitted. To be more precise, here σ appears instead of $\sqrt{\sigma}$ from (3.8).

(iii) Inequality (3.9) exhibits a better order than the bound, which can be derived from (3.5) and (3.6).

If, in Theorem 3.2, we set $p = \overline{p}$ and $s = 1$, we obtain the results regarding the compound binomial approximation.

COROLLARY 3.2. *Let the assumptions of Corollary 3.1 be valid. If $F \in \mathcal{S}$ is concentrated on the set $\mathbb{Z}\setminus\{0\}$, then*

$$\begin{aligned} \|\mathrm{GPB}(n,\,\mathbf{p},\,F) - \mathrm{Bi}(n,\,\overline{p},\,F)\| &\leq C\,\sqrt{\sigma}\,\frac{\gamma_2}{\lambda^2}, \\ |\mathrm{GPB}(n,\,\mathbf{p},\,F) - \mathrm{Bi}(n,\,\overline{p},\,F)|_h &\leq C\,\lfloor h+1\rfloor\,\frac{\gamma_2}{\lambda^{5/2}}, \end{aligned} \tag{3.10}$$

where, for (3.10)*, we assume that F has finite variance σ^2.*

In the previous results, the method of proof does not allow us to get reasonable estimates of absolute constants. However, in the special case, when F is a symmetric distribution concentrated on two points we are able to obtain asymptotic sharp constants.

THEOREM 3.3. *Let $\alpha \in (0,\infty)$ and $F = 2^{-1}(I_\alpha + I_{-\alpha})$. Let*

$$c_2^{(1)} = 0.35007\ldots, \qquad c_2^{(2)} = 0.06882\ldots, \qquad c_2^{(3)} = 0.14960\ldots$$

be defined as in Lemma 4.7 below. If the conditions in (3.7) *are satisfied, then*

$$\Big| \|\mathrm{GPB}(n,\,\mathbf{p},\,F)-\mathrm{Bi}(n,\,\overline{p},\,F)\| - c_2^{(1)}\,\frac{\gamma_2}{\lambda^2}\Big| \le C\frac{\gamma_2}{\lambda^{5/2}}, \tag{3.11}$$

$$\Big| |\mathrm{GPB}(n,\,\mathbf{p},\,F)-\mathrm{Bi}(n,\,\overline{p},\,F)| - c_2^{(2)}\,\frac{\gamma_2}{\lambda^2}\Big| \le C\,\frac{\gamma_2}{\lambda^{5/2}}, \tag{3.12}$$

$$\Big| |\mathrm{GPB}(n,\,\mathbf{p},\,F)-\mathrm{Bi}(n,\,\overline{p},\,F)|_0 - c_2^{(3)}\,\frac{\gamma_2}{\lambda^{5/2}}\Big| \le C\,\frac{\gamma_2}{\lambda^3}. \tag{3.13}$$

Remark 3. (i) In view of (3.8), one may ask why, in (3.11), the variance $\sigma^2=\alpha^2$ of F does not occur. The answer is simply that, similar to (2.10), we may assume that $\alpha=1$.

(ii) From (3.11), it follows that, under the assumptions of Theorem 3.3,

$$\|\mathrm{GPB}(n,\,\mathbf{p},\,F)-\mathrm{Bi}(n,\,\overline{p},\,F)\| \sim c_2^{(1)}\,\frac{\gamma_2}{\lambda^2}, \tag{3.14}$$

as $\lambda\to\infty$. Here, as usual, $\sim$ means that the quotient of both sides tends to one. In particular, (3.14) is valid if we assume that $\overline{p}\le 0.3$, $\theta\to 0$, and $\lambda\to\infty$. Similar relations hold for the Kolmogorov and local norms.

4. Proofs

4.1 *Norm estimates*

For several proofs below, we need the following well-known relations

$$\|VW\|\le\|V\|\,\|W\|,\qquad |VW|\le|V|\,\|W\|,\qquad |VW|_h\le|V|_h\,\|W\|,$$
$$|W|\le\|W\|,\qquad |W|_h\le\|W\|, \tag{4.1}$$

where $V,W\in\mathcal{M}$, $h\in[0,\,\infty)$. Note that, if $W(\mathbb{R})=0$, then $\max\{|W|,\,|W|_h\}\le 2^{-1}\|W\|$. As usual, for $m\in\mathbb{Z}_+$ and complex valued $x\in\mathbb{C}$, let $\binom{x}{m}=\prod_{k=1}^m[(x-k+1)/k]$.

LEMMA 4.1. *If* $F\in\mathcal{F}$, $t\in(0,\infty)$, $j\in\mathbb{N}$, $n\in\mathbb{Z}_+$, $p\in(0,1)$, *then*

$$\|(F-I)^2\exp\{t(F-I)\}\|\le\frac{3}{te}, \tag{4.2}$$

$$\|(F-I)^j\exp\{t(F-I)\}\|\le\frac{\sqrt{j!}}{t^{j/2}}, \tag{4.3}$$

$$\|(F-I)^j(qI+pF)^n\|\le\binom{n+j}{j}^{-1/2}(pq)^{-j/2} \tag{4.4}$$

$$\le\sqrt{e}\,j^{1/4}\Big(\frac{n}{n+j}\Big)^{n/2}\Big(\frac{j}{(n+j)pq}\Big)^{j/2}. \tag{4.5}$$

For a proof of (4.2) and (4.3), see Roos (2001b, Lemma 3) and Roos (2003, Lemma 4), respectively. For (4.4) and (4.5), see Roos (2000, Lemma 4).

Lemma 4.2. *Let $F \in \mathcal{F}$, $n \in \mathbb{N}$, $p \in [0,1]$, $r \in (-\infty, 1)$, $t = rnp$, and*

$$g(p) = \sum_{k=2}^{\infty} \frac{(2p)^{k-2}}{k!}(k-1) = \frac{e^{2p}(e^{-2p} - 1 + 2p)}{(2p)^2}.$$

If $p\, g(p) < (1-r)/2$, then

$$\sup_{F \in \mathcal{F}} \|(I + p(F-I))^n \exp\{-t(F-I)\}\| \leq \Big(1 - \frac{2p\, g(p)}{1-r}\Big)^{-1}.$$

Proof. Let $F \in \mathcal{F}$ and $y = np - t = (1-r)np > 0$. Then

$$\begin{aligned} T &:= \|(I + p(F-I))^n \exp\{-t(F-I)\}\| \\ &= \|[(I + p(F-I)) \exp\{-p(F-I)\}]^n \exp\{y(F-I)\}\| \\ &= \|[I + p^2(F-I)^2 R]^n \exp\{y(F-I)\}\|, \end{aligned}$$

with

$$R = \sum_{k=2}^{\infty} \frac{(-p)^{k-2}(1-k)}{k!}(F-I)^{k-2}, \qquad \|R\| \leq g(p).$$

Therefore, by using (4.3),

$$\begin{aligned} T &\leq \sum_{j=0}^{n} \binom{n}{j} p^{2j} \|R\|^j \Big\|(F-I)^j \exp\Big\{\frac{y}{2}(F-I)\Big\}\Big\|^2 \\ &\leq \sum_{j=0}^{n} \frac{n!}{(n-j)!\, y^j} (2p^2\, g(p))^j \leq \Big(1 - \frac{2p\, g(p)}{1-r}\Big)^{-1}. \end{aligned}$$

The lemma is proved. □

Lemma 4.3. *Let $G \in \mathcal{F}$, $t \in (0, \infty)$, and $j \in \mathbb{N}$. Then*

(4.6) $$\inf_{u \in \mathbb{R}} |(I_u G - I)^j \exp\{t(I_u G - I)\}| \leq \frac{C(j)}{t^{j/2 + j/(2j+2)}}.$$

The estimate (4.6) was proved in Čekanavičius (1995, Theorem 3.1).

For the results with respect to symmetric distributions, we need the following result.

Lemma 4.4. *Let $F \in \mathcal{S}$, $t \in (0, \infty)$, $j \in \mathbb{N}$, and $h \in [0, \infty)$. Then*

(4.7) $$|(F-I)^j \exp\{t(F-I)\}| \leq \frac{C(j)}{t^j},$$

(4.8) $$|(F-I)^j \exp\{t(F-I)\}|_h \leq \frac{C(j)}{t^j} \widetilde{Q}_h^{1/(2j+1)} (|\ln \widetilde{Q}_h| + 1)^{6j(j+1)/(2j+1)},$$

where $\widetilde{Q}_h := \widetilde{Q}_{h,t,F} := |\exp\{4^{-1}t(F-I)\}|_h$.

PROOF. For $h > 0$, inequality (4.8) follows from the more general Theorem 1.1 in Čekanavičius (1995). Note that, in this theorem, there is a misprint in the power of the last factor (compare the statement of the theorem in the paper with its Equation (4.25)). For $h = 0$, (4.8) is valid as well, since, for $W \in \mathcal{M}$ and $h \in [0, \infty)$, $|W|_h \leq \liminf_{r\downarrow h} |W|_r$ and $h \mapsto \widetilde{Q}_h$ is continuous from the right; see Hengartner and Theodorescu (1973, Theorem 1.1.4). The estimate (4.7) follows from (4.8). □

In what follows, we need the Fourier transform $\widehat{W}(x) = \int_{\mathbb{R}} e^{ixy}\, dW(y)$, $(x \in \mathbb{R})$ of a finite signed measure $W \in \mathcal{M}$. Here, i denotes the complex unit. It is easy to check that, for $V, W \in \mathcal{M}$ and $a, x \in \mathbb{R}$,

$$\widehat{\exp\{W\}}(x) = \exp\{\widehat{W}(x)\}, \quad \widehat{VW}(x) = \widehat{V}(x)\widehat{W}(x), \quad \widehat{I_a}(x) = e^{ixa}, \quad \widehat{I}(x) = 1.$$

LEMMA 4.5. *Let $W \in \mathcal{M}$ be concentrated on $\mathbb{Z}$ satisfying $\sum_{k\in\mathbb{Z}} |k||W(\{k\})| < \infty$. Then, for all $a \in \mathbb{R}$ and $b \in (0, \infty)$,*

$$\|W\|^2 \leq \frac{1 + b\pi}{2\pi} \int_{-\pi}^{\pi} \Big(|\widehat{W}(x)|^2 + \frac{1}{b^2}\Big|\frac{d}{dx}(e^{-ixa}\widehat{W}(x))\Big|^2\Big)\, dx. \tag{4.9}$$

Further,

$$|W|_0 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |\widehat{W}(x)|\, dx. \tag{4.10}$$

PROOF. The proof of (4.9) can be found, for example, in Presman (1985, Lemma on page 419). As was pointed out by Presman, this inequality would be equivalent to a corresponding lemma by Esseen in the lattice case (cf. Ibragimov and Linnik (1971, page 29)). Inequality (4.10) is an immediate consequence of the well-known inversion formula $W(\{k\}) = (2\pi)^{-1} \int_{-\pi}^{\pi} \widehat{W}(x) e^{-ikx}\, dx$, $(k \in \mathbb{Z})$. □

LEMMA 4.6. *Let $j \in \mathbb{Z}_+$ and $t \in (0, \infty)$. If $F \in \mathcal{S}$ is concentrated on the set $\mathbb{Z} \setminus \{0\}$, then*

$$\|(F - I)^j \exp\{t(F - I)\}\| \leq 3.6\, j^{1/4}\sqrt{1 + \sigma}\, \Big(\frac{j}{te}\Big)^j, \qquad (j \neq 0), \tag{4.11}$$

$$|(F - I)^j \exp\{t(F - I)\}|_0 \leq 2\Big(\frac{j + 1/2}{te}\Big)^{j+1/2}, \tag{4.12}$$

where, for (4.11), we assume that F has finite variance σ^2. If $F = 2^{-1}(I_\alpha + I_{-\alpha})$ for some $\alpha \in (0, \infty)$, then

$$\|(F - I)^j \exp\{t(F - I)\}\| \leq \frac{j!}{t^j}. \tag{4.13}$$

PROOF. The proof of (4.13) is easily done by using (4.3) and the fact that, under the present assumptions,

$$F-I=-\frac{1}{2}(I_{-\alpha}-I)(I_\alpha-I). \tag{4.14}$$

We now prove (4.12) by using (4.10). Let $F\in\mathcal{S}$ be concentrated on $\mathbb{Z}\setminus\{0\}$ and $W=(F-I)^j\exp\{t(F-I)\}$. Then, for $x\in\mathbb{R}$,

$$\begin{aligned}\widehat{F}(x)&=2\sum_{k=1}^{\infty}F(\{k\})\cos(kx),\qquad 1-\widehat{F}(x)=4\sum_{k=1}^{\infty}F(\{k\})\sin^2\Big(\frac{kx}{2}\Big)\ge 0,\\ \widehat{W}(x)&=(\widehat{F}(x)-1)^j\exp\{t(\widehat{F}(x)-1)\}.\end{aligned}$$

It is easy to check that, for each $G\in\mathcal{F}$ concentrated on $\mathbb{Z}$, we have

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|\widehat{G}(x)|^2\,dx=\sum_{k=-\infty}^{\infty}(G(\{k\}))^2\le|G|_0.$$

Therefore, for arbitrary $A>0$, applying (3.6), we obtain

$$\int_{-\pi}^{\pi}\exp\{A(\widehat{F}(x)-1)\}\,dx\le 2\pi\Big|\exp\Big\{\frac{A}{2}(F-I)\Big\}\Big|_0\le 2\pi\sqrt{\frac{2}{Ae}}. \tag{4.15}$$

Hence, using (4.10),

$$\begin{aligned}|W|_0&\le\frac{1}{2\pi}\int_{-\pi}^{\pi}(1-\widehat{F}(x))^j\exp\{t(\widehat{F}(x)-1)\}\,dx\\ &=\frac{1}{2\pi}\int_{-\pi}^{\pi}(1-\widehat{F}(x))^j\exp\Big\{\frac{tj}{j+1/2}(\widehat{F}(x)-1)\Big\}\exp\Big\{\frac{t}{2j+1}(\widehat{F}(x)-1)\Big\}\,dx\\ &\le\frac{1}{2\pi}\sup_{x\ge0}\Big(x^j\exp\Big\{-\frac{tjx}{j+1/2}\Big\}\Big)\int_{-\pi}^{\pi}\exp\Big\{\frac{t}{2j+1}(\widehat{F}(x)-1)\Big\}\,dx\\ &\le 2\Big(\frac{j+1/2}{te}\Big)^{j+1/2}.\end{aligned}$$

Inequality (4.12) is shown. We now prove (4.11) by using (4.9) with $a=0$. The parameter b will be chosen later. By the same arguments as above, we derive

$$\int_{-\pi}^{\pi}|\widehat{W}(x)|^2\,dx\le 4\pi\Big(\frac{2j+1/2}{2te}\Big)^{2j+1/2}. \tag{4.16}$$

Note that

$$\begin{aligned}\Big|\frac{d}{dx}\widehat{F}(x)\Big|&=4\Big|\sum_{k=1}^{\infty}kF(\{k\})\sin\Big(\frac{kx}{2}\Big)\cos\Big(\frac{kx}{2}\Big)\Big|\\ &\le 4\Big(\sum_{k=1}^{\infty}k^2F(\{k\})\sum_{j=1}^{\infty}F(\{j\})\sin^2\Big(\frac{jx}{2}\Big)\Big)^{1/2}\le\sqrt{2}\sigma(1-\widehat{F}(x))^{1/2}.\end{aligned}$$

Hence, letting $B=1-0.18j^{-1}$, we obtain

$$\begin{aligned}\Big|\frac{d}{dx}\widehat{W}(x)\Big|^2&\le\Big|\frac{d}{dx}\widehat{F}(x)\Big|^2(1-\widehat{F}(x))^{2j-2}|j-t(1-\widehat{F}(x))|^2\exp\{2t(\widehat{F}(x)-1)\}\\ &\le 2jt\sigma^2\Big(\frac{j}{t}\Big)^{2j}\sup_{y\ge0}(y^{2j-1}(1-y)^2\exp\{-2jBy\})\exp\{2t(1-B)(\widehat{F}(x)-1)\}.\end{aligned}$$

By elementary calculus, we see that the sup-term is bounded by $0.692/(je^{2j})$. With the help of (4.15), we therefore get

$$\int_{-\pi}^{\pi}\Big|\frac{d}{dx}\widehat{W}(x)\Big|^2\,dx \le \frac{173}{125}\,\sigma^2 t\Big(\frac{j}{te}\Big)^{2j}\int_{-\pi}^{\pi}\exp\Big\{\frac{9\,t}{25\,j}(\widehat{F}(x)-1)\Big\}\,dx$$

$$\le \frac{261}{40}\,\pi\sigma^2 t\Big(\frac{j}{te}\Big)^{2j+1/2}. \tag{4.17}$$

If $t \le 0.108$, then, since $\sigma \ge 1$,

$$\|W\|^2 \le 4^j \le 2^{2j-1}(1+\sigma)\Big(\frac{0.108}{t}\Big)^{2j} \le (3.6)^2\sqrt{j}(1+\sigma)\Big(\frac{j}{te}\Big)^{2j}.$$

If $t \ge 0.108$, then letting

$$b = \frac{\sigma\sqrt{t}}{\pi\sqrt{0.108}}$$

and taking into account (4.16) and (4.17), we obtain

$$\|W\|^2 \le \frac{1+b\pi}{2\pi}\Big[4\pi\Big(\frac{j+1/4}{te}\Big)^{2j+1/2} + \frac{261}{40}\,\frac{\pi\sigma^2\,t}{b^2}\Big(\frac{j}{te}\Big)^{2j+1/2}\Big]$$
$$\le (3.6)^2\sqrt{j}\,(1+\sigma)\Big(\frac{j}{te}\Big)^{2j},$$

where we used the simple inequality $(1+1/(4j))^{2j+1/2} \le (5/4)^{5/2}$, $(j \in \mathbb{N})$. The proof of (4.11) is completed. □

Remark 4. By using Tsaregradskii's (1958) inequality, it would be possible to derive an estimate for $|(F-I)^j\exp\{t(F-I)\}|$, $(j \in \mathbb{N})$ under the same assumptions as for (4.11). However, the resulting bound would have been of worse order than the one in (4.11); see also Remark 2(ii).

4.2 *Asymptotically sharp norm estimates*

In what follows, let $H_j(z)$ be the Hermite polynomial of degree $j \in \mathbb{Z}_+$, satisfying, for $z \in \mathbb{R}$,

$$H_j(z) \;=\; j!\sum_{m=0}^{\lfloor j/2\rfloor}\frac{(-1)^m(2z)^{j-2m}}{(j-2m)!\,m!},$$

$$h_j(z) := \frac{1}{\sqrt{2\pi}}\frac{d^j}{dz^j}e^{-z^2/2} = \frac{(-1)^j\,e^{-z^2/2}}{2^{(j+1)/2}\sqrt{\pi}}\,H_j\Big(\frac{z}{\sqrt{2}}\Big). \tag{4.18}$$

LEMMA 4.7. *Let $j \in \mathbb{Z}_+$, $t, \alpha \in (0,\infty)$, and $F = 2^{-1}(I_\alpha + I_{-\alpha})$. Then*

$$\Big|\|(F-I)^j\exp\{t(F-I)\}\| - \frac{2c_j^{(1)}}{t^j}\Big| \le \frac{C(j)}{t^{j+1/2}}, \qquad (j \ne 0), \tag{4.19}$$

$$\Big||(F-I)^j\exp\{t(F-I)\}| - \frac{2c_j^{(2)}}{t^j}\Big| \le \frac{C(j)}{t^{j+1/2}}, \qquad (j \ne 0), \tag{4.20}$$

$$\Big||(F-I)^j\exp\{t(F-I)\}|_0 - \frac{2c_j^{(3)}}{t^{j+1/2}}\Big| \le \frac{C(j)}{t^{j+1}}, \tag{4.21}$$

where

$$c_j^{(1)} = \frac{1}{2^{j+1}} \int_{\mathbb{R}} |h_{2j}(x)|\, dx,$$

(4.22)
$$c_j^{(2)} = \frac{1}{2^{j+1}} \sup_{x\in\mathbb{R}} |h_{2j-1}(x)|,$$

$$c_j^{(3)} = \frac{1}{2^{j+1}} \sup_{x\in\mathbb{R}} |h_{2j}(x)|.$$

In particular, we have

$$c_2^{(1)} = \frac{1}{2}\sqrt{\frac{3}{\pi}} \exp\Big\{-\frac{3-\sqrt{6}}{2}\Big\}\Big(\sqrt{3-\sqrt{6}} + e^{-\sqrt{6}}\sqrt{3+\sqrt{6}}\Big), \tag{4.23}$$

$$c_2^{(2)} = \frac{1}{8}\sqrt{\frac{3}{\pi}} \exp\Big\{-\frac{3-\sqrt{6}}{2}\Big\}\sqrt{3-\sqrt{6}}, \tag{4.24}$$

$$c_2^{(3)} = \frac{3}{8\sqrt{2\pi}}. \tag{4.25}$$

The constants given in (4.23)–(4.25) are important for Theorem 3.3. For the proof of Lemma 4.7, we need the following result, which is a slight but trivial improvement of Proposition 3 in Roos (1999).

LEMMA 4.8. *Let $j \in \mathbb{Z}_+$, S be a set, and $b : (0,\infty)\times\mathbb{R}\times S \longrightarrow \mathbb{R}$ be a bounded function. Then, for $t \in (0,\, \infty)$,*

$$\sup_{x\in S}\sup_{z\in\mathbb{R}}\Big[(1+z^2)\Big|t^{(j+1)/2}\Delta^j \mathrm{po}(\lfloor t + z\sqrt{t} + b(t,\, z,\, x)\rfloor,\, t) - (-1)^j h_j(z)\Big|\Big] \leq \frac{C(j)}{\sqrt{t}},$$

where $\Delta^0\mathrm{po}(\cdot,\, t) = \mathrm{po}(\cdot,\, t)$ denotes the counting density of the Poisson distribution with mean t and, for $m \in \mathbb{Z}$ and $j \in \mathbb{N}$,

$$\Delta^j\mathrm{po}(m,\, t) = \Delta^{j-1}\mathrm{po}(m-1,\, t) - \Delta^{j-1}\mathrm{po}(m,\, t).$$

PROOF OF LEMMA 4.7. We may assume that $t \geq 1$. In view of (4.14) and the simple relation

$$(F-I)^j \exp\{t(F-I)\} = \sum_{m=0}^{\infty} \Delta^j\mathrm{po}(m,\, t)\, F^m, \qquad (F \in \mathcal{F})$$

(cf. Roos (1999, Lemma 1)), we see that, letting $t_0 = t/2$,

$$\begin{aligned}(F-I)^j \exp\{t(F-I)\} &= \frac{1}{(-2)^j}(I_{-\alpha}-I)^j \exp\Big\{\frac{t}{2}(I_{-\alpha}-I)\Big\}(I_\alpha - I)^j \exp\Big\{\frac{t}{2}(I_\alpha - I)\Big\} \\ &= \frac{1}{(-2)^j}\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty} \Delta^j\mathrm{po}(m_1,\, t_0)\, \Delta^j\mathrm{po}(m_2,\, t_0) I_{(m_2-m_1)\alpha} \\ &= \frac{1}{(-2)^j}\sum_{k\in\mathbb{Z}}\sum_{m=0}^{\infty} \Delta^j\mathrm{po}(m,\, t_0)\, \Delta^j\mathrm{po}(k+m,\, t_0) I_{k\alpha}.\end{aligned}$$

This gives

$$T^{(1)} := \|(F-I)^j \exp\{t(F-I)\}\| = \frac{1}{2^j}\sum_{k\in\mathbb{Z}}\Big|\sum_{m=0}^{\infty}\Delta^j\mathrm{po}(m,\,t_0)\,\Delta^j\mathrm{po}(k+m,\,t_0)\Big|.$$

By using simple transformations, we obtain

$$T^{(1)} = \frac{\sqrt{t_0}}{2^j}\int_{\mathbb{R}}\Big|\int_{\mathbb{R}}\Delta^j\mathrm{po}(\lfloor t_0+y_1\sqrt{t_0}\rfloor,\,t_0)\Delta^j\mathrm{po}(\lfloor t_0+z\sqrt{t_0}+b_0\rfloor,\,t_0)\,dy_1\Big|dy_2,$$

where $b_0=\lfloor y_2\rfloor - y_2\in(-1,\,0]$ and $z=y_1+y_2/\sqrt{t_0}$. From Lemma 4.8, it follows that

$$\Delta^j\mathrm{po}(\lfloor t_0+y_1\sqrt{t_0}\rfloor,\,t_0) = \frac{(-1)^j h_j(y_1)}{t_0^{(j+1)/2}} + \frac{b_1}{(1+y_1^2)t_0^{(j+2)/2}},$$
$$\Delta^j\mathrm{po}(\lfloor t_0+z\sqrt{t_0}+b_0\rfloor,\,t_0) = \frac{(-1)^j h_j(z)}{t_0^{(j+1)/2}} + \frac{b_2}{(1+z^2)t_0^{(j+2)/2}},$$

where b_1 and b_2 are functions of $(j,\,t_0,\,y_1)$ and $(j,\,t_0,\,z,b_0)$, resp., with $\max\{|b_1|,\,|b_2|\}\le C(j)$. Combining this with the above, we arrive at

$$T^{(1)} = \frac{\sqrt{t_0}}{2^j}(R_0+R_1),$$

where

$$R_0 = \int_{\mathbb{R}}\Big|\int_{\mathbb{R}}\frac{h_j(y_1)\,h_j(z)}{t_0^{j+1}}\,dy_1\Big|dy_2 = \int_{\mathbb{R}}\Big|\int_{\mathbb{R}}\frac{h_j(y_1)\,h_j(y_1+y_2)}{t_0^{j+1/2}}\,dy_1\Big|dy_2$$

and R_1 is a quantity satisfying

$$\begin{aligned}|R_1| &\le \int_{\mathbb{R}}\int_{\mathbb{R}}\frac{|b_2|\,|h_j(y_1)|}{(1+z^2)t_0^{j+3/2}}\,dy_1dy_2 + \int_{\mathbb{R}}\int_{\mathbb{R}}\frac{|b_1|\,|h_j(z)|}{(1+y_1^2)t_0^{j+3/2}}\,dy_1dy_2\\ &\quad+\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{|b_1\,b_2|}{(1+y_1^2)(1+z^2)t_0^{j+2}}\,dy_1dy_2\\ &= R_2+R_3+R_4, \qquad \text{say.}\end{aligned}$$

It is easily shown that

$$R_2\le\frac{C(j)}{t_0^{j+1}},\qquad R_3\le\frac{C(j)}{t_0^{j+1}},\qquad R_4\le\frac{C(j)}{t_0^{j+3/2}},$$

giving $|R_1|\le C(j)t_0^{-(j+1)}$, since $t_0=t/2\ge 1/2$. This implies that

$$\Big|T^{(1)}-\frac{2c_j^{(1)}}{t^j}\Big| \le C(j)\sqrt{t_0}\,|R_1| \le \frac{C(j)}{t^{j+1/2}},$$

where

$$c_j^{(1)} = \frac{1}{2}\int_{\mathbb{R}}\Big|\int_{\mathbb{R}}h_j(y_1)\,h_j(y_1+y_2)\,dy_1\Big|dy_2.$$

Similarly,

$$
\begin{aligned}
T^{(2)} &:= |(F-I)^j \exp\{t(F-I)\}| \\
&= \frac{1}{2^j} \sup_{k\in\mathbb{Z}} \Big| \sum_{m=0}^{\infty} \Delta^j \mathrm{po}(m,\, t_0)\, \Delta^{j-1}\mathrm{po}(k+m,\, t_0) \Big| \\
&= \frac{\sqrt{t_0}}{2^j} \sup_{y_2\in\mathbb{R}} \Big| \int_{\mathbb{R}} \Delta^j \mathrm{po}(\lfloor t_0 + y_1\sqrt{t_0}\rfloor,\, t_0) \Delta^{j-1}\mathrm{po}(\lfloor t_0 + z\sqrt{t_0} + b_0\rfloor,\, t_0)\, dy_1 \Big|, \\
T^{(3)} &:= |(F-I)^j \exp\{t(F-I)\}|_0 \\
&= \frac{1}{2^j} \sup_{k\in\mathbb{Z}} \Big| \sum_{m=0}^{\infty} \Delta^j \mathrm{po}(m,\, t_0)\, \Delta^{j}\mathrm{po}(k+m,\, t_0) \Big| \\
&= \frac{\sqrt{t_0}}{2^j} \sup_{y_2\in\mathbb{R}} \Big| \int_{\mathbb{R}} \Delta^j \mathrm{po}(\lfloor t_0 + y_1\sqrt{t_0}\rfloor,\, t_0) \Delta^{j}\mathrm{po}(\lfloor t_0 + z\sqrt{t_0} + b_0\rfloor,\, t_0)\, dy_1 \Big|,
\end{aligned}
$$

and

$$
\Big| T^{(2)} - \frac{2c_j^{(2)}}{t^j} \Big| \le \frac{C(j)}{t^{j+1/2}}, \qquad \Big| T^{(3)} - \frac{2c_j^{(3)}}{t^{j+1/2}} \Big| \le \frac{C(j)}{t^{j+1}},
$$

where

$$
c_j^{(2)} = \frac{1}{2} \sup_{y_2\in\mathbb{R}} \Big| \int_{\mathbb{R}} h_j(y_1)\, h_{j-1}(y_1+y_2)\, dy_1 \Big|, \qquad c_j^{(3)} = \frac{1}{\sqrt{2}} \sup_{y_2\in\mathbb{R}} \Big| \int_{\mathbb{R}} h_j(y_1)\, h_j(y_1+y_2)\, dy_1 \Big|.
$$

By partial integration, (4.18), and some simple substitutions, we see that, for $y_2 \in \mathbb{R}$, $k, \ell \in \mathbb{Z}_+$ with $k \ge \ell$, and $m = k + \ell$,

$$
\begin{aligned}
\int_{\mathbb{R}} h_k(y_1)\, h_\ell(y_1+y_2)\, dy_1 &= (-1)^\ell \int_{\mathbb{R}} h_m(y_1)\, h_0(y_1+y_2)\, dy_1 \\
&= \frac{(-1)^k\, e^{-y_2^2/4}}{2^{m/2+1}\pi} \int_{\mathbb{R}} e^{-x^2} H_m\Big(\frac{x}{\sqrt{2}} - \frac{y_2}{2^{3/2}}\Big)\, dx.
\end{aligned}
$$

Using the well-known summation theorem for the Hermite polynomials

$$
H_m(x_1+x_2) = \frac{1}{2^{m/2}} \sum_{r=0}^{m} \binom{m}{r} H_{m-r}(x_1\sqrt{2})\, H_r(x_2\sqrt{2}), \qquad (x_1, x_2 \in \mathbb{R}),
$$

the orthogonality relation

$$
\int_{\mathbb{R}} e^{-x^2} H_{r_1}(x)\, H_{r_2}(x)\, dx = \begin{cases} \sqrt{\pi}\, 2^{r_1}\, r_1!, & \text{if } r_1 = r_2, \\ 0, & \text{if } r_1 \ne r_2, \end{cases} \qquad (r_1,\, r_2 \in \mathbb{Z}_+)
$$

(see Szegö (1975, formulas (5.5.1) and (5.5.11), pages 105–106)), and (4.18), we obtain

$$
\int_{\mathbb{R}} h_k(y_1)\, h_\ell(y_1+y_2)\, dy_1 = \frac{(-1)^k\, e^{-y_2^2/4}}{2^{m+1}\sqrt{\pi}} H_m\Big(-\frac{y_2}{2}\Big) = \frac{(-1)^\ell}{2^{(m+1)/2}}\, h_m\Big(-\frac{y_2}{\sqrt{2}}\Big).
$$

From the above, the equalities in (4.22) follow. The identities (4.23)–(4.25) are shown by using straight forward calculus. This completes the proof of the lemma. □

Lemma 4.9. *Let $j \in \mathbb{Z}_+$, $n \in \mathbb{N}$, $0 < p \le C < 1/2$, $\alpha \in (0, \infty)$, $F = 2^{-1}(I_\alpha + I_{-\alpha})$, and $c_j^{(1)}$, $c_j^{(2)}$, and $c_j^{(3)}$ be defined as in Lemma 4.7. Then*

$$\Big| \|(F-I)^j(I+p(F-I))^n\| - \frac{2c_j^{(1)}}{(np)^j}\Big| \le \frac{C(j)}{(np)^{j+1/2}}, \qquad (j \neq 0), \tag{4.26}$$

$$\Big| |(F-I)^j(I+p(F-I))^n| - \frac{2c_j^{(2)}}{(np)^j}\Big| \le \frac{C(j)}{(np)^{j+1/2}}, \qquad (j \neq 0), \tag{4.27}$$

$$\Big| |(F-I)^j(I+p(F-I))^n|_0 - \frac{2c_j^{(3)}}{(np)^{j+1/2}}\Big| \le \frac{C(j)}{(np)^{j+1}}. \tag{4.28}$$

Proof. We may assume that $np \ge 1$. We have

$$\|(F-I)^j(I+p(F-I))^n\| = \|(F-I)^j \exp\{np(F-I)\}\| + R_1^{(1)},$$

where

$$\begin{aligned} |R_1^{(1)}| &\le \|(F-I)^j[(I+p(F-I))^n - \exp\{np(F-I)\}]\| \\ &= \|([(I+p(F-I))\exp\{-p(F-I)\}]^n - I)(F-I)^j \exp\{np(F-I)\}\| \\ &= \|([I+R_2]^n - I)(F-I)^j \exp\{np(F-I)\}\| \end{aligned}$$

and

$$\begin{aligned} R_2 &= (I+p(F-I))\exp\{-p(F-I)\} - I \\ &= -p^2(F-I)^2 \sum_{k=2}^{\infty} \frac{(-p(F-I))^{k-2}}{k!}(k-1). \end{aligned}$$

Therefore, letting $g(p)$ be defined as in Lemma 4.2,

$$|R_1^{(1)}| \le \sum_{r=1}^{n} \binom{n}{r} (p^2 g(p))^r \|(F-I)^{j+2r} \exp\{np(F-I)\}\|.$$

The latter norm term can be estimated by using (4.13). In fact, for $r \in \mathbb{N}$,

$$\begin{aligned} \|(F-I)^{j+2r} \exp\{np(F-I)\}\| &\le \|(F-I)^{j+2} \exp\{\beta np(F-I)\}\| \, \|F-I\|^{r-1} \\ &\quad \times \|(F-I)^{r-1} \exp\{(1-\beta)np(F-I)\}\| \\ &\le \frac{(j+2)!\, 2^{r-1}\, (r-1)!}{(\beta np)^{j+2}((1-\beta)np)^{r-1}}, \end{aligned}$$

where $\beta \in (0, 1)$ is arbitrary. This gives

$$|R_1^{(1)}| \le \frac{(j+2)!\, p\, g(p)}{\beta^{j+2}(np)^{j+1}} \sum_{r=1}^{\infty} \Big(\frac{2p\, g(p)}{1-\beta}\Big)^{r-1}.$$

Since $g(1/2) = 1$ and $p \le C < 1/2$, we can choose a suitable $\beta \in (0, 1)$ such that

$$\frac{2p\, g(p)}{1-\beta} \le C < 1.$$

Hence $|R_1^{(1)}| \leq C(j)p(np)^{-(j+1)}$. By (4.19), we obtain

$$\Big| \|(F-I)^j (I+p(F-I))^n\| - \frac{2c_j^{(1)}}{(np)^j} \Big| \leq |R_1^{(1)}| + \frac{C(j)}{(np)^{j+1/2}} \leq \frac{C(j)}{(np)^{j+1/2}}.$$

Hence, (4.26) follows. Similarly, by using (4.12),

$$\begin{aligned} |(F-I)^j (I+p(F-I))^n| &= |(F-I)^j \exp\{np(F-I)\}| + R_1^{(2)}, \\ |(F-I)^j (I+p(F-I))^n|_0 &= |(F-I)^j \exp\{np(F-I)\}|_0 + R_1^{(3)}, \end{aligned}$$

where $|R_1^{(2)}| \leq C(j)p(np)^{-(j+1)}$ and $|R_1^{(3)}| \leq C(j)p(np)^{-(j+3/2)}$. By (4.20) and (4.21), we obtain (4.27) and (4.28). The lemma is shown. □

4.3 *Proofs of the theorems*

Proof of Theorems 3.1 and 3.2. We first prove (3.3). Let $G \in \mathcal{F}$, $x = n/8$, and $y = 3/4$. By using (2.3), (4.1), and (2.11), we obtain

$$\begin{aligned} T_0 &:= \inf_{u \in \mathbb{R}} \Big| \mathrm{GPB}(n,\, \mathbf{p},\, I_u G) - \mathrm{Bi}(n,\, p,\, I_u G;\, s) \Big| \\ &= \inf_{u \in \mathbb{R}} \Big| \sum_{j=s+1}^{n} a_j(p)\, (I_u G - I)^j\, (qI + pI_u G)^{n-j} \Big| \\ &= \inf_{u \in \mathbb{R}} \Big| (I_u G - I)^{s+1} \exp\{xp(I_u G - I)\} \\ &\quad * (I + p(I_u G - I))^{n - \lfloor yn \rfloor} \exp\{-xp(I_u G - I)\} \\ &\quad * \sum_{j=s+1}^{n} a_j(p) (I_u G - I)^{j-s-1} (I + p(I_u G - I))^{(\lfloor yn \rfloor + 1) - (j+1)} \Big| \\ &\leq T_1\, T_2 \sum_{j=s+1}^{n} \frac{|a_j(p)|}{(1-2p)^{j+1}} \|(I_1 - I)^{j-s-1} (I + p(I_1 - I))^{\lfloor yn \rfloor + 1}\|, \end{aligned}$$

where $*$ denotes convolution,

$$T_1 := \inf_{u \in \mathbb{R}} \Big| (I_u G - I)^{s+1} \exp\{xp(I_u G - I)\} \Big| \leq \frac{C(s)}{(np)^{(s+1)/2 + (s+1)/(2s+4)}}$$

by Lemma 4.3, and

$$T_2 := \|(I + p(I_1 - I))^{n - \lfloor yn \rfloor} \exp\{-xp(I_1 - I)\}\| \leq \frac{1}{1 - 1.2\, g(0.3)} \leq 10.4$$

by Lemma 4.2 and the assumption $p \leq 0.3$. In Roos (2000, Lemma 1) it was shown that, for $j \in \{1, \ldots, n\}$,

$$|a_j(p)| \leq \Big(\frac{\eta(p)}{2j} \Big)^{j/2} \frac{n^{(n-j)/2}}{(n-j)^{(n-j)/2}} \leq \Big(\frac{\eta(p)\, e}{2j} \Big)^{j/2}.$$

Therefore,

$$T_0 \le \frac{C(s)(\eta(p))^{(s+1)/2}}{(np)^{(s+1)/2+(s+1)/(2s+4)}}$$
$$\times \Big[1+\sum_{j=s+2}^{n}\Big(\frac{\eta(p)\,e}{2(1-2p)^2}\Big)^{(j-s-1)/2}\frac{1}{j^{j/2}}\|(I_1-I)^{j-s-1}(I+p(I_1-I))^{\lfloor yn\rfloor+1}\|\Big],$$

where, in view of (4.5), we see that the term in brackets is bounded by

$$C\sum_{j=0}^{\infty}\Big(\frac{\eta(p)}{2(1-2p)^2(\lfloor yn\rfloor+1)pq}\Big)^{j/2}\le C.$$

Inequality (3.3) is shown. The proofs of (3.4), (3.5), (3.8), and (3.9) are quite similar to the above one. The main difference is that we have to replace T_1 with

$$T_1^{(1)} = |(F-I)^{s+1}\exp\{xp(F-I)\}| \le \frac{C(s)}{(np)^{s+1}}, \tag{4.29}$$

$$T_1^{(2)} = |(F-I)^{s+1}\exp\{xp(F-I)\}|_h \le \frac{C(s)}{(np)^{s+1}}\,Q_h^{1/(2s+3)}(|\ln Q_h|+1)^{\kappa}, \tag{4.30}$$

$$T_1^{(3)} = \|(F-I)^{s+1}\exp\{xp(F-I)\}\| \le \frac{C(s)\sqrt{\sigma}}{(np)^{s+1}}, \tag{4.31}$$

$$T_1^{(4)} = |(F-I)^{s+1}\exp\{xp(F-I)\}|_h \le \frac{C(s)\,\lfloor h+1\rfloor}{(np)^{s+3/2}}, \tag{4.32}$$

respectively, where $\kappa = 6(s+1)(s+2)/(2s+3)$. Note that it is assumed that, for (4.29) and (4.30), we have $F\in\mathcal{S}$, and that, for (4.31) and (4.32), $F\in\mathcal{S}$ is concentrated on $\mathbb{Z}\setminus\{0\}$. Further, Lemmas 4.4 and 4.6 are used. The theorem is proved. □

Proof of Theorem 3.3. We may assume that $\alpha = 1$ and that, in view of (2.6) and (2.7), $\lambda\ge 1$ and $n\ge 3$. Then, by (2.3),

$$\|\mathrm{GPB}(n,\,\mathbf{p},\,F)-(\overline{q}I+\overline{p}F)^n\| = \frac{\gamma_2}{2}\|(F-I)^2(I+\overline{p}(F-I))^n\|+R^{(1)},$$

where, by using Lemma 4.6 and Theorem 3.2,

$$\begin{aligned}|R^{(1)}| &\le \frac{\gamma_2}{2}\,\|(F-I)^2[(I+\overline{p}(F-I))^{n-2}-(I+\overline{p}(F-I))^n]\|\\ &\quad+\frac{|\gamma_3|}{3}\,\|(F-I)^3(I+\overline{p}(F-I))^{n-3}\|+\|\mathrm{GPB}(n,\,\mathbf{p},\,F)-\mathrm{Bi}(n,\,\overline{p},\,F;\,3)\|\\ &\le C\Big(\frac{\overline{p}\gamma_2}{\lambda^3}+\frac{|\gamma_3|}{\lambda^3}+\frac{\gamma_2^2}{\lambda^4}\Big)\\ &\le C\,\frac{\gamma_2}{\lambda^3}.\end{aligned}$$

The proof of (3.11) is easily completed by the help of Lemma 4.9. Inequalities (3.12) and (3.13) are similarly shown. □

Acknowledgements

This work was mostly done during the first author's stay at the Department of Mathematics at the University of Hamburg, Germany, in June/July 2003, and also the second author's stay at the Mathematical Institute of the University of Cologne, Germany, in summer term 2004. The first author would like to thank the members of the Department of Mathematics in Hamburg for their hospitality. The second author would like to thank Professor J. Steinebach for the invitation to Cologne and the Mathematical Institute for its hospitality. Finally, we thank the referees for helpful comments.

Appendix A: The Krawtchouk expansion

We now give a review of some facts from Roos (2000). In Takeuchi and Takemura (1987a), the reader can find further informations concerning the Krawtchouk expansion of the distribution of the sum of possibly dependent Bernoulli random variables. Multivariate generalizations were derived in Takeuchi and Takemura (1987b) and Roos (2001a). Let S_n be the sum of $n \in \mathbb{N}$ independent Bernoulli random variables $X_1, \dots, X_n$ with success probabilities $p_1, \dots, p_n$, that is, the distribution of S_n is given by $\mathcal{L}(S_n) = \mathrm{GPB}(n,\, \mathbf{p},\, I_1)$. Then (2.3) with $F = I_1$ is equivalent to

$$P(S_n = m) = \sum_{j=0}^{n} a_j(p)\, \Delta^j \mathrm{bi}(m,\, n-j,\, p), \tag{A.1}$$

where $m \in \mathbb{Z}_+$, $p \in [0,\, 1]$ is arbitrary,

$$\mathrm{bi}(m,\, k,\, p) = \Delta^0 \mathrm{bi}(m,\, k,\, p) = \begin{cases} \binom{k}{m} p^m\, q^{k-m}, & \text{for } m, k \in \mathbb{Z}_+,\ m \le k, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\Delta^j \mathrm{bi}(m,\, k,\, p) = \Delta^{j-1} \mathrm{bi}(m-1,\, k,\, p) - \Delta^{j-1} \mathrm{bi}(m,\, k,\, p) \qquad \text{for } j \in \mathbb{N}.$$

In fact, (A.1) and (2.3) can be derived from each other by means of the simple equality

$$\sum_{m=0}^{k+j} \Delta^j \mathrm{bi}(m,\, k,\, p)\, z^m = (1 + p(z-1))^k\, (z-1)^j$$

for $j, k \in \mathbb{Z}_+$, $z \in \mathbb{C}$. The right–hand side of (A.1), resp. of (2.3), is called the Krawtchouk expansion of $\mathcal{L}(S_n)$ with parameter p, and $a_0(p), \dots, a_n(p)$ are called the corresponding Krawtchouk coefficients. With our assumptions, relation (3.5) of Takeuchi and Takemura (1987a) is similar to (A.1). For $n, m \in \mathbb{Z}_+$ and $j \in \{0, \dots, n\}$, we have

$$\frac{d^j}{dp^j} \mathrm{bi}(m,\, n,\, p) = n^{[j]}\, \Delta^j \mathrm{bi}(m,\, n-j,\, p),$$

where $n^{[j]} = n!/(n-j)!$, and hence, in view of (A.1), we see that

$$P(S_n = m) = \sum_{j=0}^{n} \frac{a_j(p)}{n^{[j]}} \frac{d^j}{dp^j} \mathrm{bi}(m,\, n,\, p), \qquad (m \in \mathbb{Z}_+).$$

As was mentioned in Subsection 2.2, we have $a_0(p) = 1$. In what follows, we give some alternative formulae for $a_1(p), \dots, a_n(p)$. For this, we need the Krawtchouk polynomials $\mathrm{Kr}(j;\, x,\, n,\, p) \in \mathbb{R}[x]$ being orthogonal with respect to the binomial distribution. They are defined by (see Szegö (1975, formula (2.82.2), page 36))

$$\text{(A.2)} \qquad \mathrm{Kr}(j;\, x,\, n,\, p) = \sum_{k=0}^{j} \binom{n-x}{j-k} \binom{x}{k} (-p)^{j-k} q^k, \qquad (n, j \in \mathbb{Z}_+,\ x \in \mathbb{C}).$$

Then, for $j \in \{1, \dots, n\}$,

$$\text{(A.3)} \qquad a_j(p) = \sum_{1 \le k(1) < \ldots < k(j) \le n} \prod_{r=1}^{j} (p_{k(r)} - p),$$

$$\text{(A.4)} \qquad = \sum_{m=0}^{n} P(S_n = m)\, \mathrm{Kr}(j;\, m,\, n,\, p)$$

$$\text{(A.5)} \qquad = \sum_{k=0}^{j} \binom{n-k}{j-k} \frac{1}{k!} (-p)^{j-k} \mu_{(k)},$$

$$\text{(A.6)} \qquad = \frac{1}{2\pi\, \alpha^j} \int_0^{2\pi} e^{-ijx} \prod_{k=1}^{n} \Big(1 + (p_k - p)\alpha e^{ix}\Big)\, dx.$$

where $\alpha \in (0, \infty)$ is arbitrary and, for $k \in \{0, \dots, n\}$,

$$\mu_{(k)} = \sum_{m=k}^{n} m^{[k]}\, P(S_n = m)$$

denotes the kth factorial moment of S_n. The equalities (2.4) and (A.3)–(A.6) can be derived with the help of the generating functions

$$\sum_{j=0}^{n} a_j(p)\, z^j = \prod_{k=1}^{n} (1 + (p_k - p)z), \qquad (z \in \mathbb{C})$$

and

$$\text{(A.7)} \qquad \sum_{j=0}^{n} \mathrm{Kr}(j;\, m,\, n,\, p)\, z^j = (1+qz)^m\, (1-pz)^{n-m},$$

for $n, m \in \mathbb{Z}_+$, $n \ge m$, and $z \in \mathbb{C}$. For (A.7), see Szegö (1975, formula (2.82.4), page 36). In view of (A.4), we see that the Krawtchouk coefficients can be defined by using the Krawtchouk polynomials. But the same is true for the differences $\Delta^j \mathrm{bi}(m,\, k,\, p)$. Indeed, for $m, k, j, r \in \mathbb{Z}_+$ and $p \in [0, 1]$, we have

$$\text{(A.8)} \qquad \binom{k+j}{j} (pq)^j\, \Delta^j \mathrm{bi}(m,\, k,\, p) = \mathrm{Kr}(j;\, m,\, k+j,\, p)\, \mathrm{bi}(m,\, k+j,\, p).$$

Using (2.4), (A.2), and (A.8), the counting densities of the signed measures $\mathrm{Bi}(n,\,p,\,I_1;\,s)$ (see (2.5)) can be derived. It turns out that, for $m \in \mathbb{Z}_+$, we have

$$\mathrm{Bi}(n,\,p,\,I_1;\,1)(\{m\}) = \mathrm{bi}(m,\,n,\,p)\Big(1 - \frac{\gamma_1(p)\,(m-np)}{npq}\Big)$$

and, if $n \in \{2,3,\ldots\}$,

$$\mathrm{Bi}(n,\,p,\,I_1;\,2)(\{m\}) = \mathrm{bi}(m,\,n,\,p)\Big(1 - \frac{\gamma_1(p)\,(m-np)}{npq} + \frac{(\gamma_1(p))^2 - \gamma_2(p)}{2\,n(n-1)\,(pq)^2}\Big[m^2 - (1+2(n-1)p)m + n(n-1)p^2\Big]\Big).$$

It is worth mentioning that, as one can expect, the first s moments of $\mathcal{L}(S_n)$ and $\mathrm{Bi}(n,\,p,\,I_1;\,s)$ coincide. This follows from the fact that, for $s \in \{0,\ldots,n\}$, $k \in \{0,\ldots,s\}$, and $\mu_{(k)}$ as above,

$$\sum_{m=k}^{n} m^{[k]}\,\mathrm{Bi}(n,\,p,\,I_1;\,s)(\{m\}) = \mu_{(k)}.$$

Finally, note that the connection between $\mathrm{Bi}(n,\,p,\,F;\,s)$, for $F \in \mathcal{F}$, and $\mathrm{Bi}(n,\,p,\,I_1;\,s)$ is given by

$$\mathrm{Bi}(n,\,p,\,F;\,s) = \sum_{m=0}^{n} \mathrm{Bi}(n,\,p,\,I_1;\,s)(\{m\})F^m.$$

References

Arak, T. V. and Zaĭtsev, A. Yu. (1988). *Uniform Limit Theorems for Sums of Independent Random Variables*, transl. from Russian, (English), Proceedings of the Steklov Institute of Mathematics, **174**(1), American Mathematical Society, Providence, RI.

Barbour, A. D., Holst, L. and Janson, S. (1992). *Poisson Approximation*, Clarendon Press, Oxford.

Barbour, A. D. and Choi, K. P. (2004). A non-uniform bound for translated Poisson approximation, *Electronic Journal of Probability*, **9**, 18–36.

Čekanavičius, V. (1995). On smoothing properties of compound Poisson distributions, *Lithuanian Mathematical Journal*, **35**, 121–135.

Čekanavičius, V. (2002). On multivariate compound distributions, *Teoriya Veroyatnostei i ee Primeneniya*, **47**, 583–594; English translation in *Theory of Probability and its Applications*, **47**, 493–506;

Čekanavičius, V. and Roos, B. (2004). Two-parametric compound binomial approximations, *Lietuvos Matematikos Rinkinys*, **44**, 443–466.

Čekanavičius, V. and Vaitkus, P. (2001). Centered Poisson approximation via Stein's method, *Lithuanian Mathematical Journal*, **41**, 319–329.

Choi, K. P. and Xia, A. (2002). Approximating the number of successes in independent trials: Binomial versus Poisson, *Annals of Applied Probability*, **12**, 1139–1148.

Ehm, W. (1991). Binomial approximation to the Poisson binomial distribution, *Statistics & Probability Letters*, **11**, 7–16.

Hengartner, W. and Theodorescu, R. (1973). *Concentration Functions*, Academic Press, New York.

Ibragimov, I. A. and Linnik, Yu. V. (1971). *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff, Groningen.

Le Cam, L. (1965). On the distribution of sums of independent random variables, In *Bernoulli–Bayes–Laplace Anniversary volume* (eds. J. Neyman and L. Le Cam), 179–202, Springer Verlag, New York.

Michel, R. (1987). An improved error bound for the compound Poisson approximation of a nearly homogeneous portfolio, *ASTIN Bulletin*, **17**, 165–169.

Presman, É. L. (1985). Approximation in variation of the distribution of a sum of independent Bernoulli variables with a Poisson law, *Teoriya Veroyatnostei i ee Primeneniya*, **30**, 391–396; English translation in *Theory of Probability and its Applications*, **30**, 417–422.

Roos, B. (1999). Asymptotics and sharp bounds in the Poisson approximation to the Poisson-binomial distribution, *Bernoulli*, **5**, 1021–1034.

Roos, B. (2000). Binomial approximation to the Poisson binomial distribution: The Krawtchouk expansion, *Teoriya Veroyatnostei i ee Primeneniya*, **45**, 328–344; see also in *Theory of Probability and its Applications*, **45,** 258–272.

Roos, B. (2001a). Multinomial and Krawtchouk approximations to the generalized multinomial distribution, *Teoriya Veroyatnostei i ee Primeneniya*, **46**, 117–133; see also in *Theory of Probability and its Applications*, **46**, 103–117.

Roos, B. (2001b). Sharp constants in the Poisson approximation, *Statistics & Probability Letters*, **52**, 155–168.

Roos, B. (2003). Improvements in the Poisson approximation of mixed Poisson distributions, *Journal of Statistical Planning and Inference*, **113**, 467–483.

Roos, B. (2004). On Hipp's compound Poisson approximations via concentration functions, to appear: *Bernoulli.*

Soon, S. Y. T. (1996). Binomial approximation for dependent indicators, *Statistica Sinica*, **6**, 703–714.

Szegö, G. (1975). *Orthogonal Polynomials*, 4th Ed., American Mathematical Society, Providence, RI.

Takeuchi, K. and Takemura, A. (1987a). On sum of 0–1 random variables. I. Univariate case, *Annals of the Institute of Statistical Mathematics*, **39**, 85–102.

Takeuchi, K. and Takemura, A. (1987b). On sum of 0–1 random variables. II. Multivariate case, *Annals of the Institute of Statistical Mathematics*, **39**, 307–324.

Tsaregradskii, I. P. (1958). On uniform approximation of the binomial distribution by infinitely divisible laws, *Teoriya Veroyatnostei i ee Primeneniya*, **3**, 470–474; English translation in *Theory of Probability and its Applications*, **3**, 434–438.