Two-parametric Compound Binomial Approximations

V. Čekanavičius* and B. Roos†

Vilnius University and University of Hamburg

Abstract

We consider two-parametric compound binomial approximation of the generalized Poisson binomial distribution. We show that the accuracy of approximation essentially depends on the symmetry or shifting of distributions and construct asymptotic expansions. For the proofs, we combine the properties of norms with the results for convolutions of symmetric and shifted distributions. In the lattice case, we use the characteristic function method. In the case of almost binomial approximation we apply Stein's method.

Keywords: compound binomial distribution, concentration function, Kolmogorov norm, shifted distributions, Stein's method, symmetric distributions, total variation norm.

*R*eceived 6 October, 2004

MSC 2000 Subject Classification: Primary 60F05; secondary 60G50.

1 Introduction

1.1 Aim of the paper

The binomial approximation is usually applied to the distribution of the sum of non-identically distributed Bernoulli random variables. In this paper, we extend the research to the case of a two-parametric compound binomial approximation to the generalized Poisson

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

†Department of Mathematics, SPST, University of Hamburg, Bundesstr. 55, 20146 Hamburg, Germany. E-mail: roos@math.uni-hamburg.de

binomial distribution. We show that the symmetry or suitable shifting of distributions significantly increase the accuracy of approximation. For the proofs, we use the properties of the exponentials of finite signed measures, relations between norms, and known results on convolutions of symmetric and shifted distributions. In the lattice case, we apply the characteristic function method. The results for the almost binomial approximation are derived using Stein's method.

1.2 Notation

Let $\mathbb{R}$ and $\mathbb{N}$ denote the sets of real numbers and positive integers, respectively. 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. $\mathcal{S}$, $\mathcal{M}$) denote the set of probability distributions (resp. symmetric probability distributions about zero, finite signed measures) on $\mathbb{R}$. The distribution concentrated at a point $u \in \mathbb{R}$ is denoted by I_u. Set $I = I_0$. All products and powers of finite signed measures are defined in the convolution sense. For $W \in \mathcal{M}$, set $W^0 = I$. Let $W = W^+ - W^-$ denote the Hahn–Jordan decomposition of W. The total variation norm, the Kolmogorov norm, and the Lévy concentration seminorm of W 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, for W concentrated on integers, $|W|_0$ is the so-called local norm of W. It is well known that, for $V, W \in \mathcal{M}$, $u \in \mathbb{R}$, and $h \in [0, \infty)$, we have

$$\|VW\| \le \|V\|\,\|W\|, \qquad |VW| \le |V|\,\|W\|, \qquad |VW|_h \le |V|_h\,\|W\|,$$
$$\|I_u V\| = \|V\|, \quad |I_u V| = |V|, \quad |I_u V|_h = |V|_h, \quad |W| \le \|W\|, \quad |W|_h \le \|W\|.$$

Note that, if $W\{\mathbb{R}\} = 0$, then, as is easily seen, $\max\{|W|,\, |W|_h\} \le 2^{-1}\|W\|$. The total variation distance between $V, W \in \mathcal{M}$ with the same total mass $V\{\mathbb{R}\} = W\{\mathbb{R}\}$ can be evaluated in terms of the total variation norm, i.e.,

$$d_{\mathrm{TV}}(V,\, W) := \sup_A |(V - W)\{A\}| = \frac{1}{2}\|V - W\|. \tag{1}$$

The supremum in (1) is taken over all Borel sets $A \subseteq \mathbb{R}$. The exponential of $W \in \mathcal{M}$ is defined by the finite signed measure

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

We denote by C positive absolute constants that may differ from line to line. Similarly, by $C(\cdot)$ we denote positive constants depending on the indicated argument only. Let

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

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

$$\lambda_k = \sum_{j=1}^{n} p_j^k, \qquad (k \in \mathbb{N}), \qquad \lambda = \lambda_1, \qquad \omega = \frac{\lambda_2}{\lambda}, \qquad \tilde{N} = \frac{\lambda^2}{\lambda_2}, \tag{2}$$

$$N = \frac{\lambda^2}{\lambda_2} - \delta, \qquad N \in \mathbb{N}, \qquad |\delta| \le \frac{1}{2}, \qquad \tilde{p} = \frac{\lambda}{N} = \frac{\omega}{1 - \delta\omega/\lambda}, \qquad \tilde{q} = 1 - \tilde{p}, \tag{3}$$

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

We always assume that $\tilde{p} \le 1$. Note that $\mathrm{Bi}(N,\, \tilde{p},\, I_1)$ and $\mathrm{GPB}(n,\, \mathbf{p},\, I_1)$ are the binomial and Poisson binomial distributions, respectively. Let S denote a random variable with distribution $\mathrm{GPB}(n,\, \mathbf{p},\, I_1)$.

The main goal of this paper is to 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 $\mathrm{Bi}(N,\, \tilde{p},\, F)$. Note that we are always interested in the case $\lambda \ge 1$, and one of our main tasks is an establishment of explicit dependence of the estimates on λ. By analogy with the Poisson approximation, we can say that we concentrate our efforts on the obtaining the 'magic factors' λ^{-r} for some $r > 0$ (see Introduction of [3]).

1.3 Known results

In general, there are two different methods for choosing the parameters of the approximating binomial distribution. The first one assumes the replacement of all p_j by their mean. For the Poisson binomial distribution, such a one-parametric approximation was considered by Ehm [9] and Roos [14]. The main disadvantage of the one-parametric approach is related to the fact that only one moment of the approximated distribution can be matched. As a typical one-parametric result, we have

$$\Big|\|\mathrm{GPB}(n,\,\mathbf{p},\,I_1)-\mathrm{Bi}(n,\,\overline{p},\,I_1)\|-\sqrt{\frac{2}{\pi\mathrm{e}}}\,\theta\Big|\leq C\,\theta\min\Big\{1,\,\theta+\frac{1}{\sqrt{\lambda(1-\overline{p})}}\Big\},\tag{4}$$

where

$$\overline{p}=\frac{\lambda}{n},\qquad\theta=\frac{\sum_{j=1}^{n}(\overline{p}-p_j)^2}{\lambda(1-\overline{p})};$$

for a slightly stronger result, see Theorem 3 in [14]. In Remark on p. 259 of [14], it is shown that $\theta\leq p_{\max}-p_{\min}$. Therefore, in view of (4), it seems that the one-parametric binomial approximation is applicable in the case where the p_j are close in some sense. Moreover, the converse holds. Indeed, from Ehm's [9] results it follows that the total variation norm term in the left-hand side of (4) is small if and only if θ is small. In what follows, whenever we speak of close p_j, we mean that θ is small. An extension of the one-parametric approach to the compound binomial case is given in [5].

The second approach is based on matching the first two moments of the Poisson binomial and binomial distributions. Thus, we have the two-parametric case. Recall that S denotes a random variable with distribution $\mathrm{GPB}(n,\,\mathbf{p},\,I_1)$. As an approximation, we use the binomial distribution $\mathrm{Bi}(N,\,\tilde{p},\,I_1)$. Obviously, both distributions have the same mean $\lambda=N\tilde{p}$. The exact matching of the variances is usually not achieved because of the necessity of $N\in\mathbb{N}$. However, the difference between the variances is small. Indeed, since $\omega/\lambda\leq 1$ and $|\delta|\leq 1/2$, we have

$$\Big|\sum_{j=1}^{n}p_jq_j-N\tilde{p}\tilde{q}\Big|=\frac{\omega^2|\delta|}{1-\delta\omega/\lambda}\leq\omega^2\leq p_{\max}^2.\tag{5}$$

As a typical two-parametric result, we have, for $0<\tilde{p}<1$,

$$\begin{aligned}&\Big\|\mathrm{GPB}(n,\,\mathbf{p},\,I_1)-\mathrm{Bi}(N,\,\tilde{p},\,I_1)\Big\|\\&\qquad\leq 2P(S\geq N+1)+\frac{2}{\tilde{q}}\Big\{\frac{\omega^2|\delta|}{\lambda-\delta\omega}+4\min\Big\{1,\,\frac{\sqrt{\mathrm{e}}}{\sqrt{\lambda-\lambda_2}}\Big\}\Big(\frac{\lambda_3}{\lambda}-\omega^2\Big)\Big\}.\end{aligned}\tag{6}$$

Estimate (6) was obtained by Čekanavičius and Vaitkus ([7], Theorem 4.1) and is an improvement of previous results of Barbour et al. ([3], p. 190) and Soon [16] if one takes into account the assumptions of this paper. Note that Soon [16] applied a two-parametric binomial approximation to a much more general case of the sum of dependent indicators. Note also that, in all three papers mentioned above, the summand $2P(S\geq N+1)$ was overlooked. However, it is often quite small. Indeed, from Bernstein's inequality it follows that

$$P(S \geq N+1) \leq \exp\Big\{-\frac{1}{4}\Big(\frac{\lambda^2}{\lambda_2} - \lambda\Big)\Big\} \leq \exp\Big\{-\frac{\lambda}{4}(1-\omega)\Big\}$$

(see [1], Theorem 1.4.1). Therefore, as a rule, the order of accuracy of approximation in (6) is determined by the second summand.

It is easy to check that if all p_j are equal, then both sides in (6) vanish. Thus, the two-parametric binomial approximation retains one of the most essential properties of the one-parametric approximation. Now, let us consider the case where the p_j are not close and uniformly bounded away from 0 and 1. Then the one-parametric binomial approximation is inaccurate, since the total variation distance between the corresponding distributions is larger than some absolute constant. On the other hand, if, in addition, $\tilde{p} \leq C < 1$, then the two-parametric binomial approximation is of accuracy $Cn^{-1/2}$. In the case mentioned, the classical Berry–Esseen inequality tells us that, for the Kolmogorov norm, one can also apply the normal approximation with accuracy $Cn^{-1/2}$. Thus, we conclude that, to some extent, the two-parametric binomial approximation combines the advantages of both the one-parametric binomial and normal approximations.

By the properties of the total variation norm (see, for example, [12]),

$$\Big\|\mathrm{GPB}(n,\, \mathbf{p},\, I_1) - \mathrm{Bi}(N,\, \tilde{p},\, I_1)\Big\| = \sup_{F \in \mathcal{F}} \Big\|\mathrm{GPB}(n,\, \mathbf{p},\, F) - \mathrm{Bi}(N,\, \tilde{p},\, F)\Big\|. \tag{7}$$

Consequently, if treated as a compound binomial approximation, estimate (6) corresponds to the worst possible case. On the other hand, if F satisfies certain structural conditions, one can use them to achieve some improvements in the accuracy. The main purpose of this paper is an investigation of the changes in the accuracy of the compound binomial approximation, when F is a symmetric or suitably shifted distribution.

As already mentioned above, the choice of parameters for the two-parametric approximation $\mathrm{Bi}(N,\, \tilde{p},\, I_1)$ is based on matching the first two moments. Recall the definition of $\tilde{N}$, ω, and S. Obviously, $\tilde{N}\omega = \lambda$ and $\tilde{N}\omega(1-\omega) = \lambda - \lambda_2$ coincide with mean and variance of S, respectively. However, we cannot take $\tilde{N}$ and ω as the binomial parameters, since, due to the definition of the binomial law, the parameter N must be a positive integer. Therefore, N is defined as an integer closest to $\tilde{N}$, and $\tilde{p}$ is chosen to satisfy $N\tilde{p} = \lambda$. Thompson [17] introduced a different approach, replacing the binomial approximation by an *almost* binomial approximation, which depends directly on $\tilde{N}$ and ω. The almost binomial distribution $\mathrm{ABi} = \mathrm{ABi}(\tilde{N}, \omega)$ is defined by

$$\mathrm{ABi}\{k\} = \begin{cases} \binom{\tilde{N}}{k}\omega^k(1-\omega)^{\tilde{N}-k}, & k \in \{0, 1, \ldots, \lfloor\tilde{N}\rfloor\}, \\ 1 - \sum_{k=0}^{\lfloor\tilde{N}\rfloor}\binom{\tilde{N}}{k}\omega^k(1-\omega)^{\tilde{N}-k}, & k = \lfloor\tilde{N}\rfloor + 1. \end{cases}$$

Here $\binom{\tilde{N}}{k} = \prod_{m=1}^{k}((\tilde{N} - m + 1)/m)$, and $\lfloor\tilde{N}\rfloor \in \mathbb{N}$ denotes the integer part of $\tilde{N}$. One may ask whether ABi is, indeed, a probability distribution. In fact, it should be clarified that $\mathrm{ABi}\{\lfloor\tilde{N}\rfloor + 1\} \geq 0$. But this easily follows from the identity

$$\sum_{k=0}^{m}\binom{x}{k}r^k(1-r)^{x-k} = 1 - x\binom{x-1}{m}\int_0^r y^m(1-y)^{x-1-m}\,\mathrm{d}y, \tag{8}$$

which holds for all $m = 0, 1, 2, \ldots$, $r \in [0,\, 1)$, and $x \in \mathbb{R}$. In the case $x \in \mathbb{N}$, this is a well-known fact (see [11], p. 110). For general x, (8) remains valid, as one can show by differentiating the left-hand side with respect to r. Thompson [17, 18] has shown that, for $A \subseteq \{0, 1, \ldots, \lfloor\tilde{N}\rfloor\}$,

$$|P(S \in A) - \mathrm{ABi}\{A\}| \leq \frac{4}{1-\omega}\left(\frac{\lambda_3}{\lambda} - \omega^2\right) + \frac{n - \lfloor\tilde{N}\rfloor - 1}{(1-\omega)(\lfloor\tilde{N}\rfloor + 1)}P(S \geq \lfloor\tilde{N}\rfloor + 2). \tag{9}$$

It should be mentioned that the bound in (9) cannot trivially be viewed as a bound for $d_{\mathrm{TV}}(\mathrm{GPB}(n,\, \mathbf{p},\, I_1),\, \mathrm{ABi})$. Indeed, the consideration of complements is not sufficient, since $\mathrm{GPB}(n,\, \mathbf{p},\, I_1)$ and ABi are concentrated on $\{0, \ldots, n\} \supseteq \{0, \ldots, \lfloor\tilde{N}\rfloor + 1\}$, respectively, and since the set A is not allowed to contain $\lfloor\tilde{N}\rfloor + 1$. Further, estimate (9) is more conservative than (6) in the sense that it is not comparable to the normal one whenever the p_j are not close. However, as observed by Thompson [18], numerical experiments show that the left-hand side of (9) seems to be of order $Cn^{-1/2}$. Therefore, one can expect that the right-hand side might be improved (see Section 2.4).

Note that there are also some other approaches, different from that of this paper (see, for example, [10]).

2 Results

2.1 General distributions

In this section, we consider shifted and symmetric distributions $F \in \mathcal{F}$. By a shifted distribution, we mean $F = I_u G$, where $G \in \mathcal{F}$ and $u \in \mathbb{R}$. Then, minimizing the norm estimate of the difference between $\mathrm{GPB}(n,\, \mathbf{p},\, I_u G)$ and $\mathrm{Bi}(N,\, \tilde{p},\, I_u G)$ with respect to u, one can expect some improvement of the accuracy of approximation. Shifted and symmetric distributions play an important rôle in compound Poisson approximations (see, for example, [1] and [12]). Note that, in the following theorem, we do not assume the finiteness of any moments.

Theorem 2.1 *Let $p_{\max} \le 1/4$ and $\lambda \ge 1$. Then we have*

$$\begin{aligned}
&\sup_{G\in\mathcal{F}} \inf_{u\in\mathbb{R}} \Big|\mathrm{GPB}(n,\, \mathbf{p},\, I_u G) - \mathrm{Bi}(N,\, \tilde{p},\, I_u G)\Big| \\
&\qquad \le \frac{C}{\lambda^{5/6}}\bigg\{\frac{1}{\lambda^{1/2}}\frac{\omega^2|\delta|}{1-\delta\omega/\lambda} + \Big(\frac{\lambda_3}{\lambda} - \omega^2\Big) + \frac{1}{\lambda^{3/2}}\sum_{j=1}^n p_j|p_j-\omega|\bigg\} && (10)\\
&\qquad \le \frac{C}{\lambda^{5/6}}, && (11)\\
&\sup_{F\in\mathcal{S}} \Big|\mathrm{GPB}(n,\, \mathbf{p},\, F) - \mathrm{Bi}(N,\, \tilde{p},\, F)\Big| \\
&\qquad \le \frac{C}{\lambda^{2}}\bigg\{\frac{\omega^2|\delta|}{1-\delta\omega/\lambda} + \Big(\frac{\lambda_3}{\lambda} - \omega^2\Big) + \frac{1}{\lambda^{2}}\sum_{j=1}^n p_j|p_j-\omega|\bigg\} && (12)\\
&\qquad \le \frac{C}{\lambda^{2}}. && (13)
\end{aligned}$$

Moreover, for all $F \in \mathcal{S}$ and $h \in (0,\infty)$, we have

$$\begin{aligned}
&\Big|\mathrm{GPB}(n,\, \mathbf{p},\, F) - \mathrm{Bi}(N,\, \tilde{p},\, F)\Big|_h \\
&\quad \le \frac{C}{\lambda^{2}}\bigg\{\frac{\omega^2|\delta|}{1-\delta\omega/\lambda} + \Big(\frac{\lambda_3}{\lambda} - \omega^2\Big) + \frac{1}{\lambda^{2}}\sum_{j=1}^n p_j|p_j-\omega|\bigg\} Q_h^{1/9}(|\ln Q_h|+1)^{40/3}. && (14)
\end{aligned}$$

Here Q_h denotes the Lévy concentration function

$$Q_h = Q_{h,\lambda,F} = \Big|\exp\Big\{\frac{\lambda}{20}(F-I)\Big\}\Big|_h.$$

Remark 2.1 (a) Estimates (11) and (13) are added for demonstration of the accuracy when p_j are not close. In this sense, estimates (10) and (12) are sharper than (11) and (13), since, if $p_j = p$ for all j, they are equal to zero. Note that, in contrast to (10)–(13), the bound in (14) depends on F.

(b) The assumption $p_{\max} \le 1/4$ is a consequence of the method of proof.

2.2 Lattice distributions

It seems that, for the total variation norm, estimates of the same accuracy and generality as in (10)–(13) are unobtainable. However, for a lattice distribution F with finite second moment, some analogue of (12) holds. Moreover, in this case, we obtain the estimates for the Kolmogorov norm and concentration seminorm with explicit constants.

Theorem 2.2 *Let $p_{\max} \le 1/4$ and $\lambda \ge 1$. Let $F \in \mathcal{S}$ be concentrated on $\{\pm 1, \pm 2, \ldots\}$. Then, for $h \in [0, \infty)$, we have*

$$\begin{aligned}
\Big|\mathrm{GPB}(n,\,\mathbf{p},\,F) - \mathrm{Bi}(N,\,\tilde{p},\,F)\Big|_h &\le \frac{\lfloor h+1 \rfloor}{\lambda^{5/2}}\bigg\{\frac{\omega^2|\delta|}{1-\delta\omega/\lambda}\Big(0.82 + \frac{0.87}{\lambda} + \frac{3.69}{\lambda^2}\Big) \\
&\quad + 1.62\Big(\frac{\lambda_3}{\lambda} - \omega^2\Big) + \frac{2.42}{\lambda^2}\sum_{j=1}^n p_j|p_j - \omega|\bigg\} \qquad (15)\\
&\le C\,\frac{h+1}{\lambda^{5/2}}.
\end{aligned}$$

If, in addition, F has finite variance σ^2, then

$$\begin{aligned}
\Big\|\mathrm{GPB}(n,\,\mathbf{p},\,F) - \mathrm{Bi}(N,\,\tilde{p},\,F)\Big\| &\le C\frac{\sqrt{\sigma}}{\lambda^2}\bigg\{\frac{\omega^2|\delta|}{1-\delta\omega/\lambda} + \Big(\frac{\lambda_3}{\lambda} - \omega^2\Big) \\
&\quad + \frac{1}{\lambda^2}\sum_{j=1}^n p_j|p_j - \omega|\bigg\} \qquad (16)\\
&\le C\frac{\sqrt{\sigma}}{\lambda^2},\\
\Big|\mathrm{GPB}(n,\,\mathbf{p},\,F) - \mathrm{Bi}(N,\,\tilde{p},\,F)\Big| &\le \frac{\sigma}{\lambda^2}\bigg\{\frac{\omega^2|\delta|}{1-\delta\omega/\lambda}\Big(0.39 + \frac{0.34}{\lambda} + \frac{1.27}{\lambda^2}\Big) \\
&\quad + 0.64\Big(\frac{\lambda_3}{\lambda} - \omega^2\Big) + \frac{0.83}{\lambda^2}\sum_{j=1}^n p_j|p_j - \omega|\bigg\} \qquad (17)\\
&\le C\,\frac{\sigma}{\lambda^2}.
\end{aligned}$$

Remark 2.2 (a) Taking $h < 1$, from (15) we get an estimate for the local norm, which is of better order than one that can be obtained from the more general estimate (14) (see also [5]). It also contains explicit constants.

(b) The main advantage of estimate (17), in comparison to (16) and (12), is its explicit constants. Note that, unlike estimate (12), it contains the undesirable factor σ, which cannot be less than one. Estimate (17) is obtained by a direct application of Tsaregradskii's inequality (see [19] or (39) below) and reveals, to some extent, the limitations of such an approach.

2.3 Asymptotic expansions

In this section, we introduce asymptotic expansions for the two-parametric compound binomial approximation. As above, we investigate the cases of shifted and symmetric distributions F. However, we begin with the general case $F \in \mathcal{F}$. In this section, we assume that $\lambda \geq 1$ and $p_{\max} \leq 1/4$, which implies that $0 < \tilde{p} \leq 2/7$ (see (46) and (47) below). As explained in the introduction, we concentrate ourselves on the 'magic factors.' The construction of an asymptotic expansion can, therefore, be outlined as follows. Under the assumptions above, the accuracy in (6) is at least of order $C\lambda^{-1/2}$. The task is to choose a finite signed measure which, when added to the binomial approximation, leads to a remainder of order $C\lambda^{-1}$. The addition of the next member of the asymptotic expansion gives a remainder of order $C\lambda^{-3/2}$, and so on.

Let us begin with the simple identity

$$\prod_{j=1}^{n}(q_jI + p_jF) = (\tilde{q}I + \tilde{p}F)^N \exp\Big\{\sum_{k=1}^{\infty}(a_k(F) + b_k(F))\Big\}, \quad F \in \mathcal{F}, \tag{18}$$

where

$$\begin{aligned} a_k(F) &= \frac{(-1)^{k+1}}{k}\lambda(\omega^{k-1} - \tilde{p}^{k-1})(F-I)^k, \\ b_k(F) &= \frac{(-1)^{k+1}}{k}\sum_{j=1}^{n}p_j(p_j^{k-1} - \omega^{k-1})(F-I)^k, \qquad k \in \mathbb{N}. \end{aligned}$$

Note that $a_1(F)$, $b_1(F)$, and $b_2(F)$ are zero measures. In what follows, we have to expand the exponential of signed measure in some power series and then to collect the summands having the same order of smallness with respect to λ. The order of smallness will be determined by the magnitude of the total variation norm of their convolution with the compound binomial distribution. The measure $a_k(F)$ contains a factor which is bounded by unity. Indeed,

$$\frac{\lambda}{k}|\omega^{k-1} - \tilde{p}^{k-1}| \leq \lambda|\omega - \tilde{p}| = \frac{\omega^2|\delta|}{1 - \delta\omega/\lambda} \leq \omega^2 \leq 1,$$

(see (5)). Consequently, by the properties of the total variation norm (see (7)), for $k \in \mathbb{N}$, we have

$$
\begin{aligned}
\sup_{F\in\mathcal{F}} \|a_k(F)(\tilde{q}I+\tilde{p}F)^N\| &\le \sup_{F\in\mathcal{F}} \|(F-I)^k(\tilde{q}I+\tilde{p}F)^N\| \\
&= \|(I_1-I)^k(\tilde{q}I+\tilde{p}I_1)^N\| \\
&\le C(k)\,\frac{1}{\lambda^{k/2}}. \qquad (19)
\end{aligned}
$$

Similarly,

$$
\sup_{F\in\mathcal{F}} \|b_k(F)(\tilde{q}I+\tilde{p}F)^N\| \le \lambda \sup_{F\in\mathcal{F}} \|(F-I)^k(\tilde{q}I+\tilde{p}F)^N\| \le C(k)\,\frac{1}{\lambda^{(k-2)/2}}. \qquad (20)
$$

The estimates for $F = I_1$ can be obtained, for example, by applying Lemma 4 from [14]. As follows from (19) and (20), the estimate of the norm of the convolution of $a_k(F)$ with the compound binomial distribution is comparable to the similar estimate for $b_{k+2}(F)$. Therefore, we use the following formal expansion in powers of x:

$$
\exp\Big\{\sum_{k=1}^{\infty}(a_k(F)+b_{k+2}(F))x^k\Big\} = I + A_1(F)x + A_2(F)x^2 + \dots\ .
$$

Taking into account (19) and (20), one can prove that

$$
\sup_{F\in\mathcal{F}} \Big\|A_k(F)(\tilde{q}I+\tilde{p}F)^N\Big\| \le C(k)\frac{1}{\lambda^{k/2}}, \qquad k\in\mathbb{N}. \qquad (21)
$$

Thus, as an approximation, we propose to use the finite signed measure

$$
(\tilde{q}I+\tilde{p}F)^N\Big(I+\sum_{k=1}^{s}A_k(F)\Big), \qquad s\in\{0,1,\dots\}. \qquad (22)
$$

Note that, by the formula of Faá di Bruno (see, for example, [13], pp. 135–136),

$$
\begin{aligned}
A_1(F) &= b_3(F), \qquad A_2(F) = a_2(F) + b_4(F) + \frac{1}{2}b_3^2(F), \\
A_k(F) &= \sum\nolimits^{*} \prod_{m=1}^{k} \frac{1}{l_m!}(a_m(F)+b_{m+2}(F))^{l_m}, \qquad k\in\mathbb{N}.
\end{aligned}
$$

Here $\sum^*$ means the summation over all nonnegative integer solutions $l_1,\dots,l_k$ of the equation $l_1 + 2l_2 + \dots + kl_k = k$.

If $p_{\min}$ is bounded away from zero, then $\lambda \approx Cn$, and expansion (22) can be written as a sum of signed measures, the norms of which, due to (21), are bounded from above by powers of $n^{-1/2}$. For the case $F = I_1$ and the Kolmogorov norm, a similar result can be obtained by the Edgeworth expansion, which confirms the already noted similarity between the two-parametric binomial approximation and the normal one. Remarkably, (22) has certain advantages in comparison with the Edgeworth expansion. No continuity terms are needed for it and, unlike the Edgeworth expansion, it can be easily used for shifted distributions.

Theorem 2.3 *Let $p_{\max} \le 1/4$ and $\lambda \ge 1$. Then we have*

$$\sup_{F\in\mathcal{F}} \Big\| \mathrm{GPB}(n,\,\mathbf{p},\,F) - \mathrm{Bi}(N,\,\tilde{p},\,F)\big(I + A_1(F)\big)\Big\|$$
$$\le \frac{C}{\lambda}\bigg\{\frac{\omega^2|\delta|}{1-\delta\omega/\lambda} + \Big(\frac{\lambda_3}{\lambda} - \omega^2\Big)^2 + \frac{1}{\lambda}\sum_{j=1}^{n} p_j|p_j-\omega|\bigg\} \tag{23}$$
$$\le \frac{C}{\lambda}, \tag{24}$$
$$\sup_{G\in\mathcal{F}}\inf_{u\in\mathbb{R}} \Big| \mathrm{GPB}(n,\,\mathbf{p},\,I_uG) - \mathrm{Bi}(N,\,\tilde{p},\,I_uG)\big(I + A_1(I_uG)\big)\Big|$$
$$\le \frac{C}{\lambda^{4/3}}\bigg\{\frac{\omega^2|\delta|}{1-\delta\omega/\lambda} + \Big(\frac{\lambda_3}{\lambda} - \omega^2\Big)^2 + \frac{1}{\lambda}\sum_{j=1}^{n} p_j|p_j-\omega|\bigg\} \tag{25}$$
$$\le \frac{C}{\lambda^{4/3}}. \tag{26}$$

Comparing (6) with (11) and (24) with (26), we see that, in both cases, shifting adds a 'magic factor' $\lambda^{-1/3}$.

Let us consider the symmetric case $F \in \mathcal{S}$. We use the same principles of construction as above. However, we switch to the Kolmogorov norm and use the following estimates:

$$\sup_{F\in\mathcal{S}} \Big| a_k(F)(\tilde{q}I + \tilde{p}F)^N \Big| \le C(k)\,\frac{1}{\lambda^k}, \tag{27}$$
$$\sup_{F\in\mathcal{S}} \Big| b_k(F)(\tilde{q}I + \tilde{p}F)^N \Big| \le C(k)\,\frac{1}{\lambda^{k-1}}, \qquad k\in\mathbb{N}, \tag{28}$$

the proof of which can be found below (see (61) and (62)). Here, as follows from (27) and (28), the estimate of the norm of the convolution of $a_k(F)$ with the compound binomial distribution is comparable to the similar estimate for $b_{k+1}(F)$. Therefore, we use the following formal expansion in powers of x:

$$\exp\Big\{\sum_{k=2}^{\infty}(a_k(F) + b_{k+1}(F))x^k\Big\} = I + \tilde{A}_2(F)x^2 + \tilde{A}_3(F)x^3 + \dots\,.$$

Taking into account (27) and (28), one can prove that

$$\sup_{F\in\mathcal{S}}\Big|\tilde{A}_k(F)(\tilde{q}I+\tilde{p}F)^N\Big| \le C(k)\,\frac{1}{\lambda^k}, \qquad k\in\mathbb{N}.$$

Thus, in the symmetric case, as an approximation, we propose to use the finite signed measure

$$(\tilde{q}I+\tilde{p}F)^N\bigg(I+\sum_{k=2}^{s}\tilde{A}_k(F)\bigg), \qquad s\in\mathbb{N}. \tag{29}$$

Note that $\tilde{A}_2(F)=a_2(F)+b_3(F)$ and that, more generally,

$$\tilde{A}_k(F)=\sum\nolimits^{**}\prod_{m=2}^{k}\frac{1}{l_m!}(a_m(F)+b_{m+1}(F))^{l_m}, \qquad k\in\{2,3,\dots\}.$$

Here $\sum^{**}$ means the summation over all nonnegative integer solutions $l_2,\dots,l_k$ of the equation $2l_2+\cdots+kl_k=k$. In what follows, we present results for the approximation by the signed measure

$$\mathrm{Bi}(N,\tilde{p}\,,F)(I+\tilde{A}_2(F))=(\tilde{q}I+\tilde{p}F)^N\bigg(I-\frac{\lambda}{2}(\omega-\tilde{p})(F-I)^2+\frac{\lambda}{3}\Big(\frac{\lambda_3}{\lambda}-\omega^2\Big)(F-I)^3\bigg).$$

Theorem 2.4 *Let $p_{\max}\le 1/4$ and $\lambda\ge 1$. Then we have*

$$\begin{aligned}\sup_{F\in\mathcal{S}}&\Big|\mathrm{GPB}(n,\,\mathbf{p},\,F)-\mathrm{Bi}(N,\,\tilde{p},\,F)(I+\tilde{A}_2(F))\Big|\\ &\le \frac{C}{\lambda^3}\bigg\{\frac{\omega^2|\delta|}{1-\delta\omega/\lambda}+\frac{1}{\lambda}\Big(\frac{\lambda_3}{\lambda}-\omega^2\Big)^2+\frac{1}{\lambda}\sum_{j=1}^{n}p_j|p_j-\omega|\bigg\} &(30)\\ &\le \frac{C}{\lambda^3}\,. &(31)\end{aligned}$$

Moreover, for all $F\in\mathcal{S}$ and $h\in(0,\infty)$, we have

$$\begin{aligned}&\Big|\mathrm{GPB}(n,\,\mathbf{p},\,F)-\mathrm{Bi}(N,\,\tilde{p},\,F)(I+\tilde{A}_2(F))\Big|_h\\ &\quad\le \frac{C}{\lambda^3}\bigg\{\frac{\omega^2|\delta|}{1-\delta\omega/\lambda}+\frac{1}{\lambda}\Big(\frac{\lambda_3}{\lambda}-\omega^2\Big)^2+\frac{1}{\lambda}\sum_{j=1}^{n}p_j|p_j-\omega|\bigg\}Q_h^{1/13}(|\ln Q_h|+1)^{252/13}. \quad (32)\end{aligned}$$

Here Q_h is the same as in Theorem 2.1.

We should note that, in the case of lattice distributions F, similar results are possible but, because of lack of space, we omit them.

2.4 Almost Binomial approximation

As noted by Thompson [18], estimate (9) does not reflect the 'normal aspect' of the two-parametric approximation correctly and, in certain situations, one should expect the approximation to be of order $O(n^{-1/2})$. This conjecture is also supported by the similarity of the almost binomial distribution to the distribution $\mathrm{Bi}(N,\,\tilde{p},\,I_1)$ and estimate (6). Below we present the proof that Thompson's conjecture is correct. We recall that the random variable S has the distribution $\mathrm{GPB}(n,\,\mathbf{p},\,I_1)$. Let $\tilde{\delta}$ denote the fractional part of $\tilde{N}$, that is $\tilde{N} = \lfloor \tilde{N} \rfloor + \tilde{\delta}$, $0 \le \tilde{\delta} < 1$.

Theorem 2.5 *If* $0 < \omega < 1$, *then*

$$\begin{aligned}\Big\|\mathrm{GPB}(n,\,\mathbf{p},\,I_1) - \mathrm{ABi}\Big\| \;\; &\le \;\; 2P(S \ge \tilde{N}+1) + \frac{2\,\omega^{\lfloor \tilde{N} \rfloor + 1}}{(1-\omega)^{1-\tilde{\delta}}} \\ &\quad + \frac{8}{(1-\omega)(1-\omega\tilde{\delta}/\lambda)}\sqrt{\frac{\mathrm{e}}{\lambda-\lambda_2}}\left(\frac{\lambda_3}{\lambda} - \omega^2\right). \qquad (33)\end{aligned}$$

This estimate indeed reflects the 'normal aspect' correctly, since, for p_j uniformly bounded away from 0 and 1/4, its accuracy is at least of order $O(n^{-1/2})$. However, it must be noted that the constants in (33) are larger than those conjectured by Thompson [18].

3 Auxiliary results

We begin with exponential smoothing estimates.

Lemma 3.1 *Let* $F \in \mathcal{F}$, $a \in (0,\infty)$, *and* $k \in \mathbb{N}$. *Then we have*

$$\|(F-I)\exp\{a(F-I)\}\| \;\le\; \sqrt{\frac{2}{a\mathrm{e}}}, \qquad (34)$$

$$\|(F-I)^2\exp\{a(F-I)\}\| \;\le\; \frac{3}{a\mathrm{e}}, \qquad (35)$$

$$\inf_{u\in\mathbb{R}} |(I_uF-I)^k\exp\{a(I_uF-I)\}| \;\le\; C(k)\,\frac{1}{a^{k/2+k/(2k+2)}}. \qquad (36)$$

Proof. For the proof of (34), (35), and (36), see [8], [15] Lemma 3, [4] Theorem 3.1, respectively. □

Lemma 3.2 *Let $F \in \mathcal{S}$, $a, h \in (0,\infty)$, and let $k \in \mathbb{N}$. Then*

$$\begin{aligned} |(F-I)^k \exp\{a(F-I)\}|_h &\leq C(k)\,\frac{1}{a^k}\,\tilde{Q}_h^{1/(2k+1)}(|\ln \tilde{Q}_h|+1)^{6k(k+1)/(2k+1)}, &(37)\\ |(F-I)^k \exp\{a(F-I)\}| &\leq C(k)\,\frac{1}{a^k}. &(38)\end{aligned}$$

Here

$$\tilde{Q}_h = \tilde{Q}_{h,a,F} = \Big|\exp\Big\{\frac{a}{4}(F-I)\Big\}\Big|_h.$$

Proof. Estimate (37) is a partial case of Theorem 1.1 from [4]. Note that, in [4], 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)). Estimate (38) follows from (37). □

In what follows, we need the Fourier transform $\widehat{W}(t)$, $t \in \mathbb{R}$ of a finite signed measure $W \in \mathcal{M}$. It is defined by $\widehat{W}(t) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tx}\, W\{dx\}$, where i denotes the complex unit. It is easy to check that, for $V, W \in \mathcal{M}$ and $a, t \in \mathbb{R}$,

$$\widehat{\exp\{W\}}(t) = \exp\{\widehat{W}(t)\}, \quad \widehat{VW}(t) = \widehat{V}(t)\widehat{W}(t), \quad \widehat{I_a}(t) = \mathrm{e}^{\mathrm{i}ta}, \quad \widehat{I}(t) = 1.$$

Note that if $W \in \mathcal{M}$ is concentrated on the integers, then the well-known Tsaregradskii [19] inequality establishes the relation between $|W|$ and $\widehat{W}(t)$ in the following way:

$$|W| \leq \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{|\widehat{W}(t)|}{|\sin(t/2)|}\,\mathrm{d}t. \tag{39}$$

The following lemma deals with the exponential smoothing inequalities for symmetric lattice distributions and was proved in [5].

Lemma 3.3 *Let $F \in \mathcal{S}$ be concentrated on $\{\pm 1, \pm 2, \ldots\}$, and let $a, v \in (0,\infty)$ and $k \in \mathbb{N}$. Then*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} (1-\widehat{F}(t))^v \exp\{a(\widehat{F}(t)-1)\}\,\mathrm{d}t \leq 2\Big(\frac{v+1/2}{a\mathrm{e}}\Big)^{v+1/2}. \tag{40}$$

If, in addition, F has finite variance σ^2, then

$$\|(F-I)^k \exp\{a(F-I)\}\| \leq 3.6k^{1/4}\sqrt{1+\sigma}\Big(\frac{k}{a\mathrm{e}}\Big)^k \leq C(k)\frac{\sqrt{\sigma}}{a^k}. \tag{41}$$

It is well known that each compound distribution can be viewed as the distribution of a random sum of independent random variables. Indeed, let $S_Y = \xi_1 + \xi_2 + \ldots + \xi_Y$, where Y is a random variable concentrated on the nonnegative integers, ξ_j has distribution F, and all these random variables are independent. Denoting the distribution of S_Y by $\varphi(F)$, we get

$$\varphi(F) = \sum_{k=0}^{\infty} P(Y = k)F^k.$$

Note that $\varphi(I_1)$ is the distribution of Y. Similarly, we define the compound distribution

$$\psi(F) = \sum_{k=0}^{\infty} P(\tilde{Y} = k)F^k.$$

Here $\tilde{Y}$ is a random variable concentrated on the nonnegative integers. If the distributions of Y and $\tilde{Y}$ are close, then one can also expect $\varphi(F)$ and $\psi(F)$ to be close. Let $\Gamma_k(\varphi)$ denote the kth factorial cumulant of Y, that is, for $z = \mathrm{e}^{\mathrm{i}t}$ with t in a neighborhood of zero,

$$\ln\Big(\sum_{k=0}^{\infty} P(Y = k)z^k\Big) = \Gamma_1(\varphi)(z-1) + \Gamma_2(\varphi)\frac{(z-1)^2}{2!} + \Gamma_3(\varphi)\frac{(z-1)^3}{3!} + \dots\ .$$

Here and henceforth, we assume that all factorial cumulants are finite. Similarly to the above, by $\Gamma_k(\psi)$, we define the kth factorial cumulant of $\tilde{Y}$. If factorial cumulants of a nonnegative integer-valued random variable behave regularly, then its distribution can be replaced by a much simpler compound Poisson law. The following lemma plays a crucial rôle here.

Lemma 3.4 *Let $\Gamma_k \in \mathbb{R}$, $k \in \mathbb{N}$ be such that, for some fixed $A \geq A_0 > 2$,*

$$|\Gamma_k| \leq \frac{(k-1)!}{A^{k-1}}\,\Gamma_1, \qquad \Gamma_1 > 0, \tag{42}$$

for all k. Set

$$f(A_0) = \frac{3}{4}A_0 \ln\Big(\frac{A_0}{A_0 - 2}\Big) - \frac{3}{2}.$$

Then, for all $t > f(A_0)$,

$$\sup_{F\in\mathcal{F}} \Big\| \exp\Big\{t\,\Gamma_1(F - I) + \sum_{k=2}^{\infty} \frac{\Gamma_k}{k!}(F - I)^k\Big\}\Big\| \leq 1 + \frac{f(A_0)}{\sqrt{2\pi}(t - f(A_0))}.$$

Note that, for example, $0.72 < f(3.5) < 0.73$.

Proof. Let $F \in \mathcal{F}$. Note that $\Gamma_1 > 0$ and, therefore, $\exp\{t\Gamma_1(F - I)\} \in \mathcal{F}$, and its total variation is equal to unity. Therefore, taking into account (35) and Stirling's formula, we obtain

$$
\begin{aligned}
&\Big\| \exp\Big\{ t\,\Gamma_1(F-I) + \sum_{k=2}^{\infty} \frac{\Gamma_k}{k!}(F-I)^k \Big\} \Big\| \\
&\qquad\le\ 1 + \sum_{r=1}^{\infty} \frac{1}{r!} \Big\| \Big(\sum_{k=2}^{\infty} \frac{\Gamma_k}{k!}(F-I)^{k-2} \Big)^r (F-I)^{2r} \exp\{t\Gamma_1(F-I)\} \Big\| \\
&\qquad\le\ 1 + \sum_{r=1}^{\infty} \frac{1}{r!} \Big(\sum_{k=2}^{\infty} \frac{|\Gamma_k|}{k!}\, 2^{k-2} \Big)^r \Big\| (F-I)^2 \exp\Big\{ \frac{t\Gamma_1}{r}(F-I) \Big\} \Big\|^r \\
&\qquad\le\ 1 + \sum_{r=1}^{\infty} \frac{(\Gamma_1 f(A_0)/3)^r}{r!} \Big(\frac{3r}{t\Gamma_1 \mathrm{e}} \Big)^r = 1 + \sum_{r=1}^{\infty} \frac{r^r}{r!} \Big(\frac{f(A_0)}{t\mathrm{e}} \Big)^r \\
&\qquad\le\ 1 + \sum_{r=1}^{\infty} \frac{1}{\sqrt{2\pi r}} \Big(\frac{f(A_0)}{t} \Big)^r \le 1 + \frac{f(A_0)}{\sqrt{2\pi}(t - f(A_0))},
\end{aligned}
$$

which completes the proof. $\square$

Remark 3.1 Conditions of type (42) are usually associated with large deviations and Poisson approximations (see, for example, [2] or [6] and the references therein).

The following proposition is one of the main tools in the subsequent proofs.

Proposition 3.1 *Let $h \in (0,\infty)$ and let, for some fixed $s \in \{2,3,\dots\}$ and $A \ge 3.5$,*

$$\Gamma_k(\varphi) = \Gamma_k(\psi) \qquad \text{for all } k \in \{1,\dots,s-1\}.$$

Suppose that

$$\max\{|\Gamma_k(\varphi)|, |\Gamma_k(\psi)|\} \le \frac{(k-1)!\,\Gamma_1(\varphi)}{A^{k-1}} \qquad \text{for all } k \in \mathbb{N} \qquad (\Gamma_1(\varphi) < \infty). \tag{43}$$

Then, for each nonnegative integer m, we have

$$
\begin{aligned}
\|\varphi(F) - \psi(F)\| \ &\le\ C \sum_{k=s}^{s+m-1} \frac{|\Gamma_k(\varphi) - \Gamma_k(\psi)|}{k!} \Big\| (F-I)^k \exp\Big\{ \frac{1}{5}\Gamma_1(\varphi)\,(F-I) \Big\} \Big\| \\
&\quad + C\,R\,\Big\| (F-I)^{s+m} \exp\Big\{ \frac{1}{5}\Gamma_1(\varphi)\,(F-I) \Big\} \Big\|,
\end{aligned} \tag{44}
$$

where

$$R = \sum_{k=s+m}^{\infty} 2^{k-s-m} \frac{|\Gamma_k(\varphi) - \Gamma_k(\psi)|}{k!} \le C(s,m) \frac{\Gamma_1(\varphi)}{A^{s+m-1}}. \tag{45}$$

Estimate (44) remains true with R given by (45) if the total variation norm $\|\cdot\|$ is everywhere replaced by the Kolmogorov norm $|\cdot|$ or by concentration seminorm $|\cdot|_h$.

Remark 3.2 If $m=0$, the first sum in the right-hand side of (44) is assumed to be zero.

Proof. Set

$$W_1=W_1(F,\varphi)=\sum_{k=2}^{\infty}\frac{\Gamma_k(\varphi)}{k!}(F-I)^k,\quad W_2=W_2(F,\psi)=\sum_{k=2}^{\infty}\frac{\Gamma_k(\psi)}{k!}(F-I)^k.$$

We shall prove (44) for the total variation norm only. The remaining bounds for the Kolmogorov and concentration norms are shown in the same way. Applying Lemma 3.4 with $A_0=3.5$, $t=4/5$, and $\Gamma_k=\tau\Gamma_k(\varphi)+(1-\tau)\Gamma_k(\psi)$, $k\in\mathbb{N}$, $\tau\in[0,1]$, we obtain

$$\begin{aligned}
\|\varphi(F)-\psi(F)\| &= \|\psi(F)(\exp\{W_1-W_2\}-I)\| \\
&= \Big\|\psi(F)\int_0^1(\exp\{\tau(W_1-W_2)\})'\,\mathrm{d}\tau\Big\| \\
&= \Big\|\psi(F)\int_0^1(W_1-W_2)\exp\{\tau(W_1-W_2)\}\,\mathrm{d}\tau\Big\| \\
&\le \Big\|(W_1-W_2)\exp\Big\{\frac{1}{5}\Gamma_1(\varphi)(F-I)\Big\}\Big\| \\
&\quad\times\int_0^1\Big\|\exp\Big\{\frac{4}{5}\Gamma_1(\varphi)(F-I)+\tau W_1+(1-\tau)W_2\Big\}\Big\|\,\mathrm{d}\tau \\
&\le C\Big\|(W_1-W_2)\exp\Big\{\frac{1}{5}\Gamma_1(\varphi)(F-I)\Big\}\Big\| \\
&\le C\sum_{k=s}^{\infty}\Big\|(F-I)^k\exp\Big\{\frac{1}{5}\Gamma_1(\varphi)(F-I)\Big\}\Big\|\frac{|\Gamma_k(\varphi)-\Gamma_k(\psi)|}{k!} \\
&\le C\sum_{k=s}^{s+m-1}\Big\|(F-I)^k\exp\Big\{\frac{1}{5}\Gamma_1(\varphi)(F-I)\Big\}\Big\|\frac{|\Gamma_k(\varphi)-\Gamma_k(\psi)|}{k!} \\
&\quad+C\Big\|(F-I)^{s+m}\exp\Big\{\frac{1}{5}\Gamma_1(\varphi)(F-I)\Big\}\Big\| \\
&\quad\times\sum_{k=s+m}^{\infty}(\|F\|+\|I\|)^{k-s-m}\frac{|\Gamma_k(\varphi)-\Gamma_k(\psi)|}{k!}.
\end{aligned}$$

Noting that $\|I\|=\|F\|=1$, we complete the proof of (44). Taking into account (43), we get

$$R\le\Gamma_1(\varphi)\sum_{k=s+m}^{\infty}\frac{2^{k-s-m+1}}{kA^{k-1}}\le C(s,m)\,\frac{\Gamma_1(\varphi)}{A^{s+m-1}}.$$

The last estimate completes the proof of the proposition. $\square$

4 Proofs

In the proofs, we constantly apply the following fact: if $\lambda \geq 1$ and $p_{\max} \leq 1/4$, then, in view of (3), $\omega \leq p_{\max} \leq 1/4$, and $|\delta| \leq 1/2$, we see that

$$\tilde{p} = \frac{\omega}{1 - \delta\omega/\lambda} \leq \frac{2}{7}\,. \tag{46}$$

Moreover,

$$\tilde{p} = \frac{\lambda}{N} \geq \frac{1}{n+1} > 0. \tag{47}$$

Proof of Theorem 2.1. We apply Proposition 3.1 with

$$\varphi(F) = \prod_{j=1}^{n}(q_j I + p_j F), \qquad \psi(F) = \big(\tilde{q}I + \tilde{p}F\big)^N \tag{48}$$

and $s = m = 2$. Then

$$\frac{\Gamma_k(\varphi)}{k!} = \frac{(-1)^{k+1}}{k}\lambda_k, \qquad \frac{\Gamma_k(\psi)}{k!} = \frac{(-1)^{k+1}}{k}N\tilde{p}^k, \qquad k \in \mathbb{N}.$$

Note that (43) is satisfied with $\Gamma_1(\varphi) = \lambda$ and $A = 3.5$. Moreover, from the definition of N and $\tilde{p}$, we get

$$\begin{aligned}
|\omega - \tilde{p}| &= \frac{\omega^2|\delta|}{\lambda - \delta\omega}, \qquad |\lambda_2 - N\tilde{p}^2| = \lambda|\omega - \tilde{p}|, \\
|\lambda_3 - N\tilde{p}^3| &= \lambda\Big|\frac{\lambda_3}{\lambda} - \tilde{p}^2\Big| \leq \lambda\Big|\frac{\lambda_3}{\lambda} - \omega^2\Big| + \lambda(\omega + \tilde{p})\,|\omega - \tilde{p}|, \\
|\lambda_k - N\tilde{p}^k| &\leq \Big|\sum_{j=1}^{n} p_j(p_j^{k-1} - \omega^{k-1})\Big| + \lambda|\omega^{k-1} - \tilde{p}^{k-1}| \\
&\leq k\sum_{j=1}^{n} p_j|p_j - \omega|\Big(\frac{1}{4}\Big)^{k-2} + k\lambda|\omega - \tilde{p}|\Big(\frac{2}{7}\Big)^{k-2}, \qquad (k \in \mathbb{N}).
\end{aligned}$$

Applying the last estimate to (45), we obtain

$$\begin{aligned}
R &= \sum_{k=4}^{\infty}\frac{2^{k-4}}{k}|\lambda_k - N\tilde{p}^k| \\
&\leq \sum_{k=4}^{\infty} 2^{k-4}\Big(\sum_{j=1}^{n} p_j|p_j - \omega|\Big(\frac{1}{4}\Big)^{k-2} + \lambda|\omega - \tilde{p}|\Big(\frac{2}{7}\Big)^{k-2}\Big) \\
&= \frac{4}{21}\lambda|\omega - \tilde{p}| + \frac{1}{8}\sum_{j=1}^{n} p_j|p_j - \omega|.
\end{aligned} \tag{49}$$

Consequently,

$$
\begin{aligned}
|\varphi(F)-\psi(F)| \;\le\; & C\Big\{\sum_{k=2}^{4}\Big|(F-I)^k\exp\Big\{\frac{1}{5}\lambda(F-I)\Big\}\Big|\,\frac{\omega^2|\delta|}{1-\delta\omega/\lambda}\\
&+\Big|(F-I)^3\exp\Big\{\frac{1}{5}\lambda(F-I)\Big\}\Big|\,\lambda\Big|\frac{\lambda_3}{\lambda}-\omega^2\Big|\\
&+\Big|(F-I)^4\exp\Big\{\frac{1}{5}\lambda(F-I)\Big\}\Big|\sum_{j=1}^{n}p_j|p_j-\omega|\Big\}. \qquad (50)
\end{aligned}
$$

To complete the proof of (12), it suffices to apply Lemma 3.2. For the proof of (14), in (50), the Kolmogorov norm $|\cdot|$ is everywhere replaced by the concentration seminorm $|\cdot|_h$. For the shifted case, in (50), we replace F by I_uG and, for $k\ge 2$, apply the estimate

$$
\begin{aligned}
\Big|(I_uG-I)^k\exp\Big\{\frac{1}{5}\lambda(I_uG-I)\Big\}\Big| \;\le\; & \Big|(I_uG-I)^2\exp\Big\{\frac{1}{10}\lambda(I_uG-I)\Big\}\Big|\\
&\times\Big\|(I_1-I)^{k-2}\exp\Big\{\frac{1}{10}\lambda(I_1-I)\Big\}\Big\|\\
\le\; & \frac{C(k)}{\lambda^{(k-2)/2}}\Big|(I_uG-I)^2\exp\Big\{\frac{1}{10}\lambda(I_uG-I)\Big\}\Big|
\end{aligned}
$$

and (36). □

Proof of Theorem 2.2. It is easy to check that, for $t\in\mathbb{R}$, we have $q_j+p_j\widehat{F}(t)\ge 1-2p_j>0$ and $\tilde q+\tilde p\widehat{F}(t)\ge 1-2\tilde p>0$. Therefore, the characteristic functions of $\mathrm{GPB}(n,\,\mathbf{p},\,F)$ and $\mathrm{Bi}(N,\,\tilde p,\,F)$ do not exceed $\exp\{\lambda(\widehat{F}(t)-1)\}$. Let us assume that $h=0$ and denote the left-hand side of (15) by T. By the inversion formula, similarly to the derivation of (50), we obtain

$$
\begin{aligned}
T \;\le\; & \frac{1}{2\pi}\int_{-\pi}^{\pi}\exp\{\lambda(\widehat{F}(t)-1)\}\Big|\sum_{j=1}^{n}\ln(1+p_j(\widehat{F}(t)-1))-N\ln(1+\tilde p(\widehat{F}(t)-1))\Big|\,\mathrm{d}t\\
\le\; & \frac{1}{2\pi}\int_{-\pi}^{\pi}\exp\{\lambda(\widehat{F}(t)-1)\}\sum_{k=2}^{\infty}\frac{|\widehat{F}(t)-1|^k}{k}|\lambda_k-N\tilde p^k|\,\mathrm{d}t\\
\le\; & \frac{1}{2\pi}\int_{-\pi}^{\pi}\exp\{\lambda(\widehat{F}(t)-1)\}\Big\{\frac{\omega^2|\delta|}{1-\delta\omega/\lambda}\Big(\frac{1}{2}|\widehat{F}(t)-1|^2+\frac{5}{28}|\widehat{F}(t)-1|^3+\frac{4}{21}|\widehat{F}(t)-1|^4\Big)\\
&+\frac{1}{3}|\widehat{F}(t)-1|^3\lambda\Big|\frac{\lambda_3}{\lambda}-\omega^2\Big|+\frac{1}{8}|\widehat{F}(t)-1|^4\sum_{j=1}^{n}p_j|p_j-\omega|\Big\}\,\mathrm{d}t.
\end{aligned}
$$

For the local estimate, it suffices to apply (40). Indeed, it can be used that $\lambda^{-1}\lambda_3-\omega^2\ge 0$ and that, under the present assumptions, $1-\widehat{F}(t)\ge 0$. Estimate (15) follows from the obvious fact that, for $W\in\mathcal{M}$ concentrated on the integers, $|W|_h\le\lfloor h+1\rfloor|W|_0$. The proof of (17) is very similar. The only difference is to apply Tsaregradskii's inequality (39) and

$$\sqrt{1-\widehat{F}(t)} \le \sigma\sqrt{2}\Big|\sin\frac{t}{2}\Big|, \qquad t\in\mathbb{R}.$$

The proof for the total variation norm follows from the analogue of (50) and from (41). □

Proof of Theorems 2.3 and 2.4. Let $\varphi(F)$ and $\psi(F)$ be defined as in (48), and let the conditions of Proposition 3.1 be satisfied. As in the proof of Theorem 2.1, we get

$$\begin{aligned}
|\Gamma_2(\varphi)-\Gamma_2(\psi)| &= \lambda|\omega-\tilde{p}| = \frac{\omega^2|\delta|}{1-\delta\omega/\lambda},\\
|\Gamma_3(\varphi)-\Gamma_3(\psi)| &\le C(\lambda|\omega-\tilde{p}|+|\lambda_3-\lambda\omega^2|),\\
\sum_{k=4}^{\infty}\frac{|\Gamma_k(\varphi)-\Gamma_k(\psi)|}{k!}2^k &\le C\bigg(\lambda|\omega-\tilde{p}|+\sum_{j=1}^{n}p_j|p_j-\omega|\bigg).
\end{aligned}$$

Set

$$\begin{aligned}
W_3 &= \frac{\Gamma_2(\varphi)-\Gamma_2(\psi)}{2}(F-I)^2+\frac{\Gamma_3(\varphi)-\Gamma_3(\psi)}{6}(F-I)^3\\
&= \frac{1}{2}\lambda(\tilde{p}-\omega)(F-I)^2+\frac{1}{3}(\lambda_3-\lambda\tilde{p}^2)(F-I)^3.
\end{aligned}$$

Then

$$\begin{aligned}
|\varphi(F)-\psi(F)(I+A_1(F))| \le{}& |\varphi(F)-\psi(F)\exp\{W_3\}|+|\psi(F)(\exp\{W_3\}-I-W_3)|\\
&+|\psi(F)(W_3-A_1(F))|. \qquad (51)
\end{aligned}$$

It is easily seen that Proposition 3.1 is applicable to $\varphi(F)$ and the finite signed measure $\psi(F)\exp\{W_3\}$ with $s=4$ and $m=0$. Using (35), we obtain

$$\begin{aligned}
|\varphi(F)-\psi(F)\exp\{W_3\}| \le{}& C\Big|(F-I)^4\exp\Big\{\frac{1}{5}\lambda(F-I)\Big\}\Big|\sum_{k=4}^{\infty}\frac{|\Gamma_k(\varphi)-\Gamma_k(\psi)|}{k!}2^k\\
\le{}& C\Big|(F-I)^2\exp\Big\{\frac{1}{10}\lambda(F-I)\Big\}\Big|\bigg(\frac{\omega^2|\delta|}{\lambda-\delta\omega}+\frac{1}{\lambda}\sum_{j=1}^{n}p_j|p_j-\omega|\bigg). \qquad (52)
\end{aligned}$$

Let W_2 be defined as in the proof of Proposition 3.1. Then, applying (34) similarly to the proof of Proposition 3.1, we get

$$\begin{aligned}
|\psi(F)(\exp\{W_3\} - I - W_3)| &\leq \Big|\psi(F)\int_0^1 (1-\tau)W_3^2\exp\{\tau W_3\}\,\mathrm{d}\tau\Big| \\
&\leq \int_0^1 (1-\tau)\Big|W_3^2\exp\{\Gamma_1(\varphi)(F-I) + W_2 + \tau W_3\}\Big|\,\mathrm{d}\tau \\
&\leq C\Big|W_3^2\exp\Big\{\frac{1}{5}\lambda(F-I)\Big\}\Big| \\
&\leq C\Big\{(\Gamma_2(\varphi)-\Gamma_2(\psi))^2\Big|(F-I)^4\exp\Big\{\frac{1}{5}\lambda(F-I)\Big\}\Big| \\
&\quad + (\Gamma_3(\varphi)-\Gamma_3(\psi))^2\Big|(F-I)^6\exp\Big\{\frac{1}{5}\lambda(F-I)\Big\}\Big| \\
&\quad + |\Gamma_2(\varphi)-\Gamma_2(\psi)||\Gamma_3(\varphi)-\Gamma_3(\psi)|\Big|(F-I)^5\exp\Big\{\frac{1}{5}\lambda(F-I)\Big\}\Big|\Big\} \\
&\leq C\Big|(F-I)^2\exp\Big\{\frac{1}{10}\lambda(F-I)\Big\}\Big|\Big\{\Big(\frac{\omega^2|\delta|}{1-\delta\omega/\lambda}\Big)^2\frac{1}{\lambda} + \Big(\frac{\lambda_3}{\lambda} - \omega^2\Big)^2\Big\}. \quad (53)
\end{aligned}$$

Finally, we have

$$|\psi(F)(W_3 - A_1(F))| \leq C\Big|(F-I)^2\exp\Big\{\frac{1}{10}\lambda(F-I)\Big\}\Big|\,\frac{\omega^2|\delta|}{1-\delta\omega/\lambda}. \quad (54)$$

Let $F = I_u G$. Then collecting estimates (51)–(54) and applying (36), we complete the proof of (25). For the proof of (23), we replace the Kolmogorov norm by the total variation norm and use (35). For the proof of (30) (resp. (32)), we replace $A_1(F)$ by $\tilde{A}_2(F)$ and use (38) (resp. (37)). □

Proof of Theorem 2.5. Let us define $\mu \in \mathcal{F}$ by

$$\mu\{j\} = \frac{1}{\tilde{C}}\binom{\tilde{N}}{j}\omega^j(1-\omega)^{\tilde{N}-j}, \quad (j \in \{0,1,\dots,\lfloor\tilde{N}\rfloor\}),$$

where

$$\tilde{C} = \sum_{j=0}^{\lfloor\tilde{N}\rfloor}\binom{\tilde{N}}{j}\omega^j(1-\omega)^{\tilde{N}-j}.$$

By (8), we see that $\tilde{C} \in (0,\, 1]$. Let us introduce the Stein operator $\mathcal{A}$ by

$$(\mathcal{A}g)(j) := \begin{cases} (\tilde{N}-j)\omega g(j+1) - j(1-\omega)g(j), & \text{if } j \in \{0,1,\dots,n\}\setminus\{\lfloor\tilde{N}\rfloor\}, \\ -\lfloor\tilde{N}\rfloor(1-\omega)g(\lfloor\tilde{N}\rfloor), & \text{if } j = \lfloor\tilde{N}\rfloor, \end{cases} \quad (55)$$

where g is a real-valued sequence defined on the nonnegative integers. For a Borel set $A \subseteq \mathbb{R}$, let $g = g_A$ be solution of the following Stein equation:

$$(\mathcal{A}g)(j) = \mathbb{I}(j \in A) - \mu\{A\}, \qquad j \in \{0,1,\dots,\lfloor\tilde{N}\rfloor\}. \quad (56)$$

Here $\mathbb{I}(j \in A) = 1$ if $j \in A$ and $\mathbb{I}(j \in A) = 0$, otherwise. We assume that $g(0) = 0$ and $g(\lfloor \tilde{N} \rfloor + 1) = g(\lfloor \tilde{N} \rfloor + 2) = ... = 0$. Stein's method strongly depends on the properties of g. Employing the results of Barbour et al. ([3], pp. 188–190), we obtain

$$\sup_{j \geq 0} |g(j+1) - g(j)| \leq \frac{1}{\lfloor \tilde{N} \rfloor \omega (1-\omega)}. \tag{57}$$

Note that our bound (57) is slightly worse than $1/(\tilde{N}\omega(1-\omega))$, which follows from [3] pp. 188–190 under the assumption that $g(0) = g(1)$ and $g(j) = g(\lfloor \tilde{N} \rfloor)$ for $j = \lfloor \tilde{N} \rfloor + 1, \lfloor \tilde{N} \rfloor + 2 \ldots$. For our case, we must take into account that, in view of (55) and (56),

$$|g(\lfloor \tilde{N} \rfloor + 1) - g(\lfloor \tilde{N} \rfloor)| = |g(\lfloor \tilde{N} \rfloor)| \leq \frac{1}{\lfloor \tilde{N} \rfloor (1-\omega)}.$$

We have

$$\begin{aligned} P(S \in A) - \mathrm{ABi}\{A\} &= \mathbb{E}(\mathcal{A}g)(S) + (\mu\{A\} - \mathrm{ABi}\{A\}) \\ &\quad + \sum_{j=\lfloor \tilde{N} \rfloor + 1}^{n} P(S = j)(\mathbb{I}(j \in A) - \mu\{A\}). \end{aligned}$$

Due to the definition of ABi, for $j \in \{0, 1, \ldots, \lfloor \tilde{N} \rfloor\}$, we have $\mathrm{ABi}\{j\} = \tilde{C}\mu(j)$. Consequently, by (8),

$$\begin{aligned} \|\mu - \mathrm{ABi}\| &= 2(1 - \tilde{C}) \\ &= 2\tilde{N} \binom{\tilde{N} - 1}{\lfloor \tilde{N} \rfloor} \int_0^{\omega} y^{\lfloor \tilde{N} \rfloor} (1-y)^{\tilde{N} - 1 - \lfloor \tilde{N} \rfloor} \, \mathrm{d}y \\ &\leq 2\tilde{N} \binom{\tilde{N} - 1}{\lfloor \tilde{N} \rfloor} \frac{1}{(1-\omega)^{1-\tilde{\delta}}} \int_0^{\omega} y^{\lfloor \tilde{N} \rfloor} \, \mathrm{d}y \\ &\leq \frac{2\,\omega^{\lfloor \tilde{N} \rfloor + 1}}{(1-\omega)^{1-\tilde{\delta}}}. \end{aligned} \tag{58}$$

Clearly, we have that

$$\Big| \sum_{j=\lfloor \tilde{N} \rfloor + 1}^{n} P(S = j)(\mathbb{I}(j \in A) - \mu\{A\}) \Big| \leq P(S \geq \tilde{N} + 1). \tag{59}$$

By Z_j, $j \in \{1, \ldots, n\}$, we denote independent Bernoulli random variables with $P(Z_j = 1) = p_j = 1 - P(Z_j = 0)$. Obviously, $S = \sum_{j=1}^{n} Z_j$. Let $S_j = S - Z_j$, and let

$$h_j = \mathbb{E}(g(S_j + 2) - g(S_j + 1)), \qquad \overline{h} = \frac{1}{\lambda} \sum_{j=1}^{n} p_j h_j.$$

We have $\tilde{N}\omega = \lambda$ and $\mathbb{E}Z_j g(S) = p_j \mathbb{E}g(S_j + 1)$. Therefore,

$$\begin{aligned}
\mathbb{E}(\mathcal{A}g)(S) &= \mathbb{E}(\lambda g(S+1) - Sg(S)) - \omega \mathbb{E}S(g(S+1) - g(S)) \\
&= \sum_{j=1}^n p_j \Big(\mathbb{E}g(S+1) - \mathbb{E}g(S_j+1) \Big) - \omega \sum_{j=1}^n p_j h_j \\
&= \sum_{j=1}^n p_j^2 h_j - \omega \sum_{j=1}^n p_j h_j = \sum_{j=1}^n p_j (p_j - \omega)(h_j - \overline{h}).
\end{aligned}$$

In [7] (see Eq. (4.11) and the subsequent remark), it was proved that

$$\Big| \sum_{j=1}^n p_j (p_j - \omega)(h_j - \overline{h}) \Big| \leq 4 \sup_j |g(j+1) - g(j)| \sqrt{\frac{\mathrm{e}}{\lambda - \lambda_2}(\lambda_3 - \lambda\omega^2)}.$$

Thus, taking into account (57), we obtain

$$\Big| \mathbb{E}(\mathcal{A}g)(S) \Big| \leq \frac{4}{(1-\omega)(1-\omega\tilde{\delta}/\lambda)} \sqrt{\frac{\mathrm{e}}{\lambda - \lambda_2}} \Big(\frac{\lambda_3}{\lambda} - \omega^2 \Big). \tag{60}$$

Now it remains to collect estimates (58), (59), and (60). Note that some of the constants in (33) were doubled because of the property (1). □

Proof of (27) and (28). Let $\Gamma_k = (k-1)!(-1)^{k+1} N \tilde{p}^k$. Note that, under our assumptions, $\tilde{p} \leq 2/7$. Therefore, taking into account Lemmas 3.4 and 3.2, similarly to the proof of (19) and (20), we obtain

$$\begin{aligned}
\sup_{F \in \mathcal{S}} &|a_k(F)(\tilde{q}I + \tilde{p}F)^N| \leq \sup_{F \in \mathcal{S}} |(F-I)^k (\tilde{q}I + \tilde{p}F)^N| \\
&= \sup_{F \in \mathcal{S}} \Big| (F-I)^k \exp\Big\{ \frac{1}{10}\lambda(F-I) + \frac{9}{10}\lambda(F-I) + \sum_{m=2}^\infty \frac{\Gamma_m}{m!}(F-I)^m \Big\} \Big| \\
&\leq \sup_{F \in \mathcal{S}} \Big| (F-I)^k \exp\Big\{ \frac{1}{10}\lambda(F-I) \Big\} \Big| \, \Big\| \exp\Big\{ \frac{9}{10}\lambda(F-I) + \sum_{m=2}^\infty \frac{\Gamma_m}{m!}(F-I)^m \Big\} \Big\| \\
&\leq C(k) \sup_{F \in \mathcal{S}} \Big| (F-I)^k \exp\Big\{ \frac{1}{10}\lambda(F-I) \Big\} \Big| \leq \frac{C(k)}{\lambda^k}
\end{aligned} \tag{61}$$

and

$$\sup_{F \in \mathcal{S}} |b_k(F)(\tilde{q}I + \tilde{p}F)^N| \leq C(k)\lambda \sup_{F \in \mathcal{S}} |(F-I)^k (\tilde{q}I + \tilde{p}F)^N| \leq \frac{C(k)}{\lambda^{k-1}}. \tag{62}$$

This completes the proof. □

Acknowledgment

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 faculty of the Department of Mathematics in Hamburg for their hospitality. The second author would like to thank Prof. J. Steinebach for the invitation to Cologne and the Mathematical Institute for its hospitality. Both authors feel indebted to Prof. P. Thompson, who kindly provided his paper on the almost binomial approximation.

References

[1] T.V. Arak and A.Yu. Zaĭtsev, *Uniform Limit Theorems for Sums of Independent Random Variables.* Transl. from Russian. *Proc. Steklov Inst. Math.*, 1988, **174**(1), AMS, Providence, RI.

[2] A. Aleškevičienė and V. Statulevičius, Large deviations in the approximation by the Poisson law, *Theory Probab. Appl.*, 2001, **46**, 589–602.

[3] A.D. Barbour, L. Holst, and S. Janson, *Poisson Approximation,* 1992, Clarendon Press, Oxford.

[4] V. Čekanavičius, On smoothing properties of compound Poisson distributions, *Lith. Math. J.*, 1995, **35**, 121–135.

[5] V. Čekanavičius and B. Roos, Compound binomial approximations, 2004, Preprint (submitted).

[6] V. Čekanavičius and P. Vaitkus, Large deviations for signed compound Poisson approximations, *Statist. Decis.*, 1997, **15**, 379-396.

[7] V. Čekanavičius and P. Vaitkus, Centered Poisson approximation via Stein's method, *Lith. Math. J.*, 2001, **41**, 319–329.

[8] P. Deheuvels and D. Pfeifer, On a relationship between Uspensky's theorem and Poisson approximations, *Ann. Inst. Statist. Math.*, 1988, **40**, 671–681.

[9] W. Ehm, Binomial approximation to the Poisson binomial distribution, *Statist. Probab. Lett.*, 1991, **11**, 7–16.

[10] Š. Jakševičius, Approximation by binomial type law, *Lith. Math. J.*, 1992, **32**, 174–180.

[11] N.L. Johnson, S. Kotz and A.W. Kemp, *Univariate Discrete Distributions,* 2nd Ed., 1992, Wiley, New York.

[12] L. Le Cam, On the distribution of sums of independent random variables. In: J. Neyman and L. Le Cam (eds), *Bernoulli, Bayes, Laplace*, Anniversary volume, Springer, Berlin, 1965, 179–202.

[13] V.V. Petrov, *Sums of Independent Random Variables,* 1975, Springer, Berlin.

[14] B. Roos, Binomial approximation to the Poisson binomial distribution: The Krawtchouk expansion, *Theory Probab. Appl.*, 2000, **45**, 258–272.

[15] B. Roos, Sharp constants in the Poisson approximation, *Statist. Probab. Lett.*, 2001, **52**, 155–168.

[16] S.Y.T. Soon, Binomial approximation for dependent indicators, *Statist. Sinica*, 1996, **6**, 703–714.

[17] P. Thompson, Almost-binomial random variables, *College Math. J.*, 2002, **33**, 235–237.

[18] P. Thompson, Almost-binomial approximation of the Poisson-binomial distribution, Unpublished manuscript, Wabash College, 2003.

[19] I.P. Tsaregradskii, On uniform approximation of the binomial distribution by infinitely divisible laws, *Theory Probab. Appl.*, 1958, **3**, 434–438.