Binomial approximation to the Markov binomial distribution

V. Čekanavičius[*] and B. Roos[†]

Vilnius University and University of Hamburg

2nd November 2006

Abstract

The Markov binomial distribution is approximated by the binomial distribution. Estimates of accuracy are obtained for the total variation and local norms. The results include second-order estimates and asymptotically sharp constants.

Keywords: Markov binomial distribution, binomial approximation, local norm, total variation norm.

MSC 2000 Subject Classification: Primary 60J10; Secondary 60F05.

1 Introduction and notation

The Markov binomial distribution is a generalization of the binomial one. Depending on the choice of parameters, both distributions can be close or even equal. To the best of our knowledge, the closeness of both distributions was not investigated in detail, though there are some general results for the binomial approximation of the sum of dependent variables, see Serfling (1975), Soon (1996), and Boutsikas and Koutras (2000). Apparently, the results of Soon (1996) and Boutsikas and Koutras (2000) cannot be applied to the Markov binomial distribution directly. Note also that numerous papers are devoted to compound Poisson approximations of the Markov binomial distribution, see Dobrushin (1953), Serfling (1975), Wang (1981), Serfozo (1986), Čekanavičius and Mikalauskas (1999), and the references therein. For papers dealing with related problems, see, for example, Campbell et al. (1994), Erhardsson (1999), and Vellaisamy (2004).

We need the following notation. Let I_k denote the distribution concentrated at an integer $k \in \mathbb{Z}$ and set $I = I_0$. Throughout this paper, we use the abbreviation

$$U = I_1 - I.$$

[*]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

In what follows, let V and W be two finite signed measures on $\mathbb{Z}$. Products and powers of V, W are understood in the convolution sense, i.e. $VW\{A\} = \sum_{k=-\infty}^{\infty} V\{A-k\}\,W\{k\}$ for a set $A \subseteq \mathbb{Z}$; further $W^0 = I$. Here and henceforth, we write $W\{k\}$ for $W\{\{k\}\}$, $(k \in \mathbb{Z})$. The total variation norm and the local norm of W are denoted by

$$\|W\| = \sum_{k=-\infty}^{\infty} |W\{k\}|, \qquad \|W\|_\infty = \sup_{k\in\mathbb{Z}} |W\{k\}|,$$

respectively. The logarithm and exponential of W are given by

$$\ln(I+W) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\, W^k \quad (\text{if } \|W\| < 1), \qquad \mathrm{e}^W = \exp\{W\} = \sum_{k=0}^{\infty} \frac{1}{k!}\, W^k.$$

Note that

$$\|VW\|_\infty \leqslant \|V\|\|W\|_\infty, \qquad \|VW\| \leqslant \|V\|\|W\|, \qquad \|\mathrm{e}^W\| \leqslant \mathrm{e}^{\|W\|}.$$

We denote by C positive absolute constants. Sometimes, to avoid possible confusion, we supply constants C with indices. The letter Θ stands for any finite signed measure on $\mathbb{Z}$ satisfying $\|\Theta\| \leqslant 1$. The values of C and Θ can vary from line to line, or even within the same line. For $x \in \mathbb{R}$ and $k \in \mathbb{N} = \{1, 2, 3, \dots\}$, we set

$$\binom{x}{k} = \frac{1}{k!}\, x(x-1)\dots(x-k+1), \qquad \binom{x}{0} = 1.$$

Let $\mathrm{Bi}(n, p^*)$ denote the binomial distribution with parameters $n \in \mathbb{N}$ and $p^* \in [0,1]$. Let $\xi_0, \xi_1, \dots, \xi_n, \dots$ be a Markov chain with the initial distribution

$$\mathrm{P}(\xi_0 = 1) = p_0, \qquad \mathrm{P}(\xi_0 = 0) = 1 - p_0, \qquad p_0 \in [0,1]$$

and transition probabilities

$$\begin{gathered}
\mathrm{P}(\xi_i = 1 \,|\, \xi_{i-1} = 1) = p, \qquad \mathrm{P}(\xi_i = 0 \,|\, \xi_{i-1} = 1) = q, \\
\mathrm{P}(\xi_i = 1 \,|\, \xi_{i-1} = 0) = \overline{q}, \qquad \mathrm{P}(\xi_i = 0 \,|\, \xi_{i-1} = 0) = \overline{p}, \\
p + q = \overline{q} + \overline{p} = 1, \qquad p, \overline{q} \in (0,1), \qquad i \in \mathbb{N}.
\end{gathered}$$

The distribution of $S_n = \xi_1 + \dots + \xi_n$ $(n \in \mathbb{N})$ is called the Markov binomial distribution. We denote it by F_n, that is $\mathrm{P}(S_n = m) = F_n\{m\}$ for $m \in \mathbb{Z}_+ = \mathbb{N} \cup \{0\}$. We should note that the definition of the Markov binomial distribution slightly varies from paper to paper, see Dobrushin (1953), Serfling (1975), and Wang (1981). Sometimes ξ_0 is added to S_n or stationarity of the chain is assumed. For example, Dobrushin (1953) assumed that $p_0 = 1$ and considered $S_{n-1} + 1$. However, if $p = \overline{q}$, then Dobrushin's Markov binomial distribution becomes a binomial distribution shifted by unity. Therefore, we use the definition above which contains the binomial distribution as a special case. Moreover, it obviously allows the rewriting of our results for $S_{n-1}+1$.

Further on, we need various characteristics of S_n. Let

$$\begin{aligned}
\nu_1 &= \frac{\overline{q}}{q+\overline{q}}, \quad \nu_2=\frac{2q\overline{q}(p-\overline{q})}{(q+\overline{q})^3}, \quad A_1=\frac{\overline{q}-p}{(q+\overline{q})^2}(\overline{q}-p_0(q+\overline{q})),\\
a &= \nu_1+\frac{A_1}{n}, \quad A_2=(\overline{q}-p)\frac{q\big(2\overline{q}+q(\overline{q}-p)\big)}{(q+\overline{q})^4}-(1-p_0)(\overline{q}-p)\frac{\overline{q}+q(\overline{q}-p)}{(q+\overline{q})^3}.
\end{aligned}$$

Note that $q+\overline{q}>0$. It is known that

$$\begin{aligned}
\mathrm{E}S_n &= n\nu_1+A_1-A_1(p-\overline{q})^n,\\
\mathrm{Var}S_n &= n(\nu_2+\nu_1-\nu_1^2)+A_1-A_1^2+2A_2\\
&\quad+(p-\overline{q})^n\Big[2nA_1\frac{\overline{q}-q}{q+\overline{q}}+A_1^2(2-(p-\overline{q})^n)-A_1-2A_2\Big],
\end{aligned}$$

see Čekanavičius and Roos (2006b). If p and $\overline{q}$ are uniformly bounded away from unity, then $(p-\overline{q})^n$ is at least of exponentially vanishing order. Therefore, na can be viewed as the main part of $\mathrm{E}S_n$.

2 Results

It is known that S_n has seven different limit laws, see Dobrushin (1953, Table 1). Typical limit distributions are the compound Poisson and the normal one. Consequently, we cannot expect that the binomial approximation is good for all values of parameters p and $\overline{q}$. However, if $p=\overline{q}$ then the Markov binomial distribution coincides with the binomial one. Therefore, our aim is to get bounds, which are equal to zero if $p=\overline{q}$.

What is known about the closeness of F_n to the binomial distribution? We formulate a consequence of a more general result of Serfling (1975, Eq. (2.4b)). For arbitrary $p^*\in(0,1)$, the estimate

$$\|F_n-\mathrm{Bi}(n,p^*)\|\leqslant 2\sum_{j=1}^n \mathrm{E}|\mathrm{P}(\xi_j=1|\xi_{j-1})-p^*| \tag{1}$$

holds. For a better understanding of (1), let us take $p^*=a$. Then from (1), we obtain

$$\|F_n-\mathrm{Bi}(n,a)\|\leqslant 2n|p-\overline{q}|\Big(1+\frac{1}{n(q+\overline{q})}\Big). \tag{2}$$

Estimate (2) may be good in the case $|p-\overline{q}|=o(n^{-1})$ only. Note that the factor n is due to the summation in (1). Our purpose is to show that, under additional assumptions, it can be replaced by a smaller factor. Let

$$p\leqslant\frac{1}{20}, \qquad \nu_1\leqslant\frac{1}{30}. \tag{3}$$

Assumption (3) was introduced by Čekanavičius and Mikalauskas (1999). Though certain smallness of p and $\overline{q}$ is required, nevertheless both parameters can be constants. Thus, even if (3) is satisfied, the limit distribution of S_n can be the normal one. On the other hand, it also allows a compound Poisson limit distribution, which occurs when $n\overline{q} \to \hat{\lambda}$ and $p \to \hat{p}$. If $\hat{p} = 0$, we have the Poisson limit distribution, see Dobrushin (1953, Table 1). Our first result is the following theorem.

Theorem 2.1 *Let condition (3) be satisfied. Then*

$$\begin{aligned} \|F_n - \mathrm{Bi}(n,a)\| &\leqslant C_1|p-\overline{q}| \min(1, p+n\overline{q}), \\ \|F_n - \mathrm{Bi}(n,a)\|_\infty &\leqslant C_2|p-\overline{q}| \min\Big(\frac{1}{\sqrt{n\overline{q}}}, p+n\overline{q}\Big). \end{aligned}$$

The right-hand side of (2) is always less or equal to $C_1|p-\overline{q}|$. Thus, in comparison to (2), we get an estimate which has no factor n.

Due to the method of proof, absolute constants in Theorem 2.1 are not given explicitly. However, we can calculate asymptotically sharp constants.

Theorem 2.2 *Let condition (3) be satisfied. Then*

$$\Big|\|F_n - \mathrm{Bi}(n,a)\| - \frac{4}{\sqrt{2\pi \mathrm{e}}}\frac{|p-\overline{q}|}{q+\overline{q}}\Big| \leqslant C_3|p-\overline{q}|\Big(|p-\overline{q}| + \frac{1}{\sqrt{n\overline{q}}}\Big), \tag{4}$$

$$\Big|\|F_n - \mathrm{Bi}(n,a)\|_\infty - \frac{1}{\sqrt{2\pi}}\frac{|p-\overline{q}|}{\sqrt{nq\overline{q}}}\Big| \leqslant C_4\frac{|p-\overline{q}|}{\sqrt{n\overline{q}}}\Big(|p-\overline{q}| + \frac{1}{\sqrt{n\overline{q}}}\Big). \tag{5}$$

Note that if, in addition, $|p-\overline{q}| = o(1)$, $n\overline{q} \to \infty$, then the right hand sides of (4) and (5) are of order $o(|p-\overline{q}|)$ and $o(|p-\overline{q}|/\sqrt{n\overline{q}})$, respectively.

The accuracy of approximation can be improved by the second-order estimates.

Theorem 2.3 *Let condition (3) be satisfied. Set* $W = \mathrm{Bi}(n,a)(I + 2^{-1}n\nu_2U^2)$. *Then*

$$\|F_n - W\| \leqslant C_5|p-\overline{q}|\Big(|p-\overline{q}| \min(n^2\overline{q}^2, 1) + (p+\overline{q}) \min\Big(1, \frac{1}{\sqrt{n\overline{q}}}\Big)\Big), \tag{6}$$

$$\|F_n - W\|_\infty \leqslant C_6|p-\overline{q}|\Big(|p-\overline{q}| \min\Big(n^2\overline{q}^2, \frac{1}{\sqrt{n\overline{q}}}\Big) + (p+\overline{q}) \min\Big(1, \frac{1}{n\overline{q}}\Big)\Big). \tag{7}$$

Note that the estimate in (6) is always less or equal to $C|p-\overline{q}|(p+\overline{q})$.

So far we considered one-parametric binomial approximation. It is possible to make use of both parameters of the binomial distribution $\mathrm{Bi}(N, \widetilde{p})$, where N and $\widetilde{p}$ are chosen in order to match two moments of S_n. The main benefit of two-parametric binomial approximation is that the estimates become comparable with the ones obtained in the normal approximation. Such two-parametric approach was used for independent and dependent indicators by Barbour et al. (1992, p. 188), Čekanavičius and Vaitkus (2001), and Soon (1996), respectively. However, in Soon's paper, N and $\widetilde{p}$ depend on the variance of independent indicators rather than on the variance of the approximated sum.

Since we want to fit two moments of the Markov binomial and binomial distributions, one should note that this is not always possible. Indeed, one can check that the Markov binomial distribution can be so close to a compound Poisson limit distribution that its second factorial cumulant becomes positive. Meanwhile the binomial distribution has negative second factorial cumulant. Therefore, we use some additional assumptions. Let $p \leqslant \overline{q}$ and $n\nu_1 \geqslant 1$. Then $\nu_2 \leqslant 0$ and $\nu_1^2 - \nu_2 \geqslant \nu_1^2 > 0$. Now we can define $N \in \mathbb{N}$ and $\widetilde{p} \in [0,1]$ in the following way:

$$N = \Big\lfloor \frac{na^2}{\nu_1^2 - \nu_2} \Big\rfloor = \frac{na^2}{\nu_1^2 - \nu_2} - \delta, \quad 0 \leqslant \delta < 1, \quad N\widetilde{p} = na.$$

From the definition of $\mathrm{Var}S_n$, it follows that the main part of the second factorial cumulant of S_n is equal to $n(\nu_2 - \nu_1^2)/2$. Now

$$\frac{1}{2}|n(\nu_2 - \nu_1^2) + N\widetilde{p}^2| = \frac{\delta(\nu_2 - \nu_1^2)^2}{2(a^2 + \delta(\nu_2 - \nu_1^2)/n)} \leqslant C\delta\overline{q}^2.$$

Thus, we see that N and $\widetilde{p}$ are indeed chosen to match two factorial cumulants (and consequently moments) of S_n closely.

Theorem 2.4 *Let condition (3) be satisfied, $p \leqslant \overline{q}$ and $n\nu_1 \geqslant 1$. Then*

$$\|F_n - \mathrm{Bi}(N,\widetilde{p})\| \leqslant C_7\Big(\sqrt{\frac{\overline{q}}{n}}(\overline{q} - p) + \frac{\delta\overline{q}}{n}\Big), \tag{8}$$

$$\|F_n - \mathrm{Bi}(N,\widetilde{p})\|_\infty \leqslant C_8\Big(\frac{\overline{q} - p}{n} + \frac{\delta}{n}\sqrt{\frac{\overline{q}}{n}}\Big). \tag{9}$$

It is clear that (8) is at least of order $O(n^{-1/2})$. Thus, in this case, it becomes comparable to the classical Berry-Esseen bound in the context of independent summands. If $p = \overline{q}$, then the right-hand sides of (8) and (9) are equal to zero. Therefore, the closeness of p and $\overline{q}$ is also reflected in the bounds.

3 Auxiliary results

In what follows, $C(k)$ denotes a positive constant depending on k.

Lemma 3.1 *Let $t \in (0,\infty)$ and $k \in \mathbb{Z}_+$. Then we have*

$$\|U^2 \mathrm{e}^{tU}\| \leqslant \frac{3}{t\mathrm{e}}, \quad \|U^k \mathrm{e}^{tU}\| \leqslant \Big(\frac{2k}{t\mathrm{e}}\Big)^{k/2}, \quad \|U^k \mathrm{e}^{tU}\|_\infty \leqslant \frac{C(k)}{t^{(k+1)/2}}. \tag{10}$$

The first inequality was proved in Roos (2001, Lemma 3). The second bound follows from formula (3.8) in Deheuvels and Pfeifer (1988) and the properties of the total variation norm. Here and throughout this paper, we set $0^0 = 1$. The third relation is a simple consequence of the formula of inversion.

Lemma 3.2 *For $n \in \mathbb{N}$ and $p^* = 1 - q^* \in (0,1)$, we have*

$$\Big| \|U^2(I+p^*U)^n\| - \frac{4}{\sqrt{2\pi \mathrm{e}}\, np^*q^*}\Big| \leqslant \frac{C}{(np^*q^*)^2}, \tag{11}$$

$$\Big| \|U^2(I+p^*U)^n\|_\infty - \frac{1}{\sqrt{2\pi}\,(np^*q^*)^{3/2}}\Big| \leqslant \frac{C}{(np^*q^*)^{5/2}}. \tag{12}$$

Lemma 3.2 was proved in Roos (2000, Lemma 8), see also Čekanavičius and Roos (2006a, Prop. 3.5 and Rem. 3.1). We now give some facts about F_n. It is known that, if condition (3) is satisfied, then F_n can be expressed as $F_n = \Lambda_1^n W_1 + \Lambda_2^n W_2$, see Čekanavičius and Mikalauskas (1999, p. 215). The properties of $\Lambda_{1,2}$ and $W_{1,2}$ are given in the following lemma.

Lemma 3.3 *Let condition (3) be satisfied. Then*

$$\Lambda_1 = I + \nu_1 U + \nu_2 U^2\Theta, \tag{13}$$

$$\Lambda_1 = I + \nu_1 U + \frac{\nu_2}{2}U^2 + C\overline{q}|p-\overline{q}|(p+\overline{q})U^3\Theta, \tag{14}$$

$$\ln\Lambda_1 = \nu_1 U + \frac{\nu_2}{2}U^2 + \sum_{j=2}^{\infty}\frac{(-1)^{j+1}}{j}\nu_1^j U^j + C\overline{q}|p-\overline{q}|(p+\overline{q})U^3\Theta, \tag{15}$$

$$\Lambda_2 = 2|p-\overline{q}|\Theta, \qquad \Lambda_2^n = C(s)\,|p-\overline{q}|^s \mathrm{e}^{-C(s)n}\Theta \quad (\textit{if } n \geqslant s \geqslant 0), \tag{16}$$

$$W_1 = I + A_1 U + C|p-\overline{q}|(p+\overline{q})U^2\Theta, \quad W_1 = I + \frac{1}{2}\Theta, \tag{17}$$

$$\ln W_1 = A_1 U + C|p-\overline{q}|(p+\overline{q})U^2\Theta, \quad W_2 = C|p-\overline{q}|U\Theta. \tag{18}$$

For any finite signed measure V on $\mathbb{Z}$ and any $t \in (0,\infty)$, we have

$$\|V\mathrm{e}^{t\ln\Lambda_1}\| \leqslant C\,\|V\mathrm{e}^{0.1t\nu_1 U}\|. \tag{19}$$

Estimate (19) *also holds if the total variation norm on both sides is replaced by the local one.*

Proof. Estimates (13), (14), (16), (17), (18), (19) can be obtained from the explicit formulas for $\Lambda_{1,2}$, $W_{1,2}$ in Čekanavičius and Mikalauskas (1999, p.p. 214–215) and are already proved in Čekanavičius and Roos (2006b). For the proof of (15) note that (13), (14) and the trivial fact $\|U\| = 2$ imply that

$$\Lambda_1 - I = \frac{3\nu_1}{2}U\Theta, \quad \Lambda_1 - I - \nu_1 U = C|p-\overline{q}|\overline{q}U^2\Theta.$$

Consequently, for $j \geqslant 2$,

$$\begin{aligned}
(\Lambda_1 - I)^j - \nu_1^j U^j &= \sum_{i=1}^{j}(\Lambda_1 - I)^{i-1}\nu_1^{j-i}U^{j-i}\big(\Lambda_1 - I - \nu_1 U\big)\\
&= C|p-\overline{q}|\overline{q}U^2\sum_{i=1}^{j}\Big(\frac{3\nu_1}{2}\Big)^{i-1}\nu_1^{j-i}U^{j-1}\Theta\\
&= Cj|p-\overline{q}|\overline{q}U^3\nu_1\Big(\frac{3\nu_1}{2}\Big)^{j-2}2^{j-2}\Theta = Cj|p-\overline{q}|\overline{q}^2U^3\Big(\frac{1}{10}\Big)^{j-2}\Theta
\end{aligned}$$

and

$$\sum_{j=2}^{\infty}\frac{(-1)^{j+1}}{j}\big((\Lambda_1-I)^j-\nu_1^jU^j\big)=C|p-\overline{q}|\overline{q}^2U^3\sum_{j=2}^{\infty}\Big(\frac{1}{10}\Big)^{j-2}\Theta=C|p-\overline{q}|\overline{q}^2U^3\Theta. \tag{20}$$

Estimate (20) combined with (14) completes the proof of (15). □

Lemma 3.4 *Let condition (3) be satisfied. Then*

$$\sum_{j=3}^{\infty}\frac{(-1)^{j+1}a^j}{j}U^j=\sum_{j=3}^{\infty}\frac{(-1)^{j+1}\nu_1^j}{j}U^j+C\frac{|p-\overline{q}|(p+\overline{q})^2}{n}U^3\Theta.$$

Proof. Note that, due to (3), we have

$$|a|\leqslant 0.12,\quad \nu_1\leqslant\frac{1}{30},\quad |a^j-\nu_1^j|\leqslant Cj\frac{|p-\overline{q}|}{n(q+\overline{q})}0.2^{j-3}(p+\overline{q})^2.$$

Now the proof is obvious. □

Lemma 3.5 *Let condition (3) be satisfied. Then, for any finite signed measure V on $\mathbb{Z}$ and any $t\in(0,\infty)$, we have*

$$\|V\exp\{t\ln(I+aU)\}\|\ \leqslant\ C\|V\exp\{0.1t\nu_1U\}\|, \tag{21}$$

$$\|V\exp\{t\ln(I+aU)+0.5t\nu_2U^2\}\|\ \leqslant\ C\|V\exp\{0.1t\nu_1U\}\|. \tag{22}$$

The estimates remain valid if the total variation norm on both sides is replaced by the local one and a is replaced by ν_1.

Proof. Noting that $a\leqslant 0.1$, we obtain

$$t\ln(I+aU)+\frac{t\nu_2}{2}U^2=taU+\frac{ta^2}{2}U^2\sum_{j=2}^{\infty}0.2^{j-2}\Theta+\frac{t\nu_1}{19}U^2\Theta=C\Theta+t\nu_1U+0.08t\nu_1U^2\Theta.$$

Moreover, applying (10), we get

$$\begin{aligned}\|\exp\{0.9t\nu_1U+0.08t\nu_1U^2\Theta\}\| &\leqslant 1+\sum_{r=1}^{\infty}\frac{1}{r!}\Big\|0.08t\nu_1U^2\exp\Big\{\frac{0.9t\nu_1}{r}U\Big\}\Big\|^r\\ &\leqslant 1+\sum_{r=1}^{\infty}\frac{\mathrm{e}^r}{r^r\sqrt{2\pi r}}\Big(\frac{0.24r}{0.90\mathrm{e}}\Big)^r\leqslant C.\end{aligned}$$

The last estimate and the properties of the norms are sufficient for the proof of (22). Estimate (21) is proved similarly. □

4 Proofs

Proof of Theorem 2.1. Let $B=I+aU$, $M_1=n\ln\Lambda_1+\ln W_1$, and $M_2=n\ln B$. We have

$$\|F_n-\mathrm{Bi}(n,a)\|\leqslant\|\Lambda_1^nW_1-B^n\|+\|\Lambda_2^n\|\|W_2\|.$$

Applying (19) and Lemma 3.5, we get

$$\begin{aligned}\|\Lambda_1^n W_1 - B^n\| &= \|\mathrm{e}^{M_1} - \mathrm{e}^{M_2}\| = \Big\|\mathrm{e}^{M_2}\int_0^1 \Big(\mathrm{e}^{(M_1-M_2)\tau}\Big)'\,\mathrm{d}\tau\Big\| \\ &\leqslant \int_0^1 \|(M_1-M_2)\mathrm{e}^{M_1\tau+M_2(1-\tau)}\|\,\mathrm{d}\tau \leqslant C\|(M_1-M_2)\mathrm{e}^{0.1n\nu_1 U}\|.\end{aligned}$$

Now it suffices to apply Lemmas 3.1, 3.3, and 3.4. The estimate for the local norm is proved similarly. □

Proof of Theorem 2.3. Let B and M_1 be defined as in the proof of Theorem 2.1. Taking into account (22) and arguing as in the proof of Theorem 2.1, we get

$$\begin{aligned}\Big\|\Lambda_1^n W_1 - B^n \exp\Big\{\frac{n\nu_2}{2}U^2\Big\}\Big\| &\leqslant C\Big\|\mathrm{e}^{0.1n\nu_1 U}\Big[M_1 - n\ln B - \frac{n\nu_2}{2}U^2\Big]\Big\| \\ &\leqslant C|p-\overline{q}|(p+\overline{q})\Big(\min\Big(n\overline{q}, \frac{1}{\sqrt{n\overline{q}}}\Big) + \min\Big(1, \frac{1}{n\overline{q}}\Big)\Big).\end{aligned}$$

Further,

$$\begin{aligned}&\Big\|B^n\Big(\exp\Big\{\frac{n\nu_2}{2}U^2\Big\} - I - \frac{n\nu_2}{2}U^2\Big)\Big\| = \frac{(n\nu_2)^2}{4}\Big\|U^2\int_0^1 B^n\exp\Big\{\tau\frac{n\nu_2}{2}U^2\Big\}(1-\tau)\,\mathrm{d}\tau\Big\| \\ &\qquad\leqslant C(n\nu_2)^2\|U^4\mathrm{e}^{0.1n\nu_1 U}\| \leqslant C\big(n\overline{q}|p-\overline{q}|\big)^2\min\Big(1, \frac{1}{(n\overline{q})^2}\Big).\end{aligned}$$

Combining the last two estimates, we get (6). The local estimate is proved using (10). □

Proof of Theorem 2.2. Due to Theorem 2.1, without loss of generality, we can assume that $n\overline{q} \geqslant 1$. Let

$$\begin{aligned}b &= \frac{4|p-\overline{q}|}{\sqrt{2\pi\mathrm{e}}\,(q+\overline{q})} = \frac{n|\nu_2|}{2}\cdot\frac{4}{\sqrt{2\pi\mathrm{e}}\,n\nu_1(1-\nu_1)}, \\ b_1 &= \frac{|p-\overline{q}|}{\sqrt{2\pi\,n q\overline{q}}} = \frac{n|\nu_2|}{2}\cdot\frac{1}{\sqrt{2\pi}\,(n\nu_1(1-\nu_1))^{3/2}}.\end{aligned}$$

Then

$$\begin{aligned}\Big|\|F_n - \mathrm{Bi}(n,a)\| - b\Big| &\leqslant \Big\|F_n - \mathrm{Bi}(n,a)\Big(I + \frac{n\nu_2}{2}U^2\Big)\Big\| \\ &\quad + \Big\|\frac{n\nu_2}{2}\,U^2\big(\mathrm{Bi}(n,\nu_1) - \mathrm{Bi}(n,a)\big)\Big\| + \Big|\Big\|\frac{n\nu_2}{2}\,U^2\mathrm{Bi}(n,\nu_1)\Big\| - b\Big|.\end{aligned}$$

Taking into account Lemma 3.5, similarly to the proof of Theorem 2.1, we get

$$\Big\|\frac{n\nu_2}{2}\,U^2\big(\mathrm{Bi}(n,\nu_1) - \mathrm{Bi}(n,a)\big)\Big\| \leqslant C\Big\|A_1 n\nu_2\,U^3\mathrm{e}^{0.1n\nu_1 U}\Big\| \leqslant C\frac{|p-\overline{q}|^2}{\sqrt{n\overline{q}}}.$$

Now the proof of (4) follows from (11) and Theorem 2.2. The estimate (5) is obtained with b replaced by b_1 and the total variation norm replaced by the local norm. □

Proof of Theorem 2.4. We give only a sketch of the proof. Due to assumption $p \leqslant \overline{q}$ we have $\nu_2 \leqslant 0$. The following estimates can be obtained:

$$\frac{\nu_1^2-\nu_2}{\nu_1} \leqslant 0.14, \quad \widetilde{p} \leqslant C(p+\overline{q}), \quad \widetilde{p} \leqslant \frac{1}{4}, \quad n\nu_1 + A_1 \geqslant \frac{18}{19},$$

$$|\widetilde{p}-a| = \Big| -\widetilde{p}\frac{\nu_2 + 2\nu_1 A_1/n + A_1^2/n^2}{\nu_1^2-\nu_2} + \frac{\widetilde{p}\delta}{n}\Big| \leqslant C\Big(|p-\overline{q}| + \frac{\delta\overline{q}}{n}\Big),$$

$$|N\widetilde{p}^j - na^j| = na\Big|\widetilde{p}^{j-1} - a^{j-1}\Big| \leqslant Cn\overline{q}^2(j-1)|\widetilde{p}-a|\Big(\frac{1}{4}\Big)^{j-3}$$
$$\leqslant Cjn\overline{q}^2\Big(\overline{q}-p+\frac{\delta\overline{q}}{n}\Big)\Big(\frac{1}{4}\Big)^{j-3}, \quad (j \geqslant 3),$$

$$\sum_{j=3}^{\infty} \frac{(-1)^{j+1}}{j} U^j \big(N\widetilde{p}^j - na^j\big) = Cn\overline{q}^2\Big(\overline{q}-p+\frac{\delta\overline{q}}{n}\Big)U^3\Theta,$$

$$n\ln\Lambda_1 + \ln W_1 - N\ln(I+\widetilde{p}U) = Cn\overline{q}^2\Big(\overline{q}-p+\frac{\delta\overline{q}}{n}\Big)U^3\Theta$$
$$+C\overline{q}(\overline{q}-p)U^2\Theta + C\delta\overline{q}^2U^2\Theta.$$

The proof of the theorem is now similar to the one of Theorem 2.1. □

Acknowledgment

This work was finished during the second author's stay at the Institute of Mathematics at Carl von Ossietzky University of Oldenburg in Summer term 2006. The second author would like to thank Professor Dietmar Pfeifer for the invitation to Oldenburg and the Institute for its hospitality. We are grateful to the referee for valuable remarks, which helped to improve the results of the paper.

References

[1] Barbour, A.D., Holst, L., Janson, S.: *Poisson Approximation.* Clarendon Press, Oxford (1992)

[2] Boutsikas, M.V., Koutras, M.V.: A bound for the distribution of the sum of discrete associated or negatively associated random variables. *Ann. Appl. Probab.* **10**, 1137–1150 (2000)

[3] Campbell, S.A., Godbole, A.P., Schaller, S.: Discriminating between sequences of Bernoulli and Markov-Bernoulli trials. *Commun. Statist. Theory Meth.* **23**, 2787–2814 (1994)

[4] Čekanavičius, V., Mikalauskas, M.: Signed Poisson approximations for Markov chains. *Stochastic. Proc. Appl.* **82**, 205–227 (1999)

[5] Čekanavičius, V., Roos, B.: An expansion in the exponent for compound binomial approximations. *Lith. Math. J.* **46**, 54–91 (2006a)

[6] Čekanavičius, V., Roos, B.: Poisson type approximations for the Markov binomial distribution. Preprint 2006-14, Faculty of Mathematics and Informatics, Vilnius University, (2006b)

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

[8] Dobrushin, R.L.: Limit theorems for a Markov chain of two states. *Izv. Akad. Nauk. USSR Ser. mat.* **17**, 291–330 (1953) (Russian). English translation in *Select. Transl. Math. Stat. and Probab.* **1**, Inst. Math. Stat. and Amer. Math. Soc. 97–134 (1961)

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

[10] Erhardsson, T.: Compound Poisson approximation for Markov chains using Stein's method. *Ann. Probab.* **27**, 565–596 (1999)

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

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

[13] Serfling, R.J.: A general Poisson approximation theorem., *Ann. Probab.* **3**, 726–731 (1975)

[14] Serfozo, R.F.: Compound Poisson approximations for sums of random variables. *Ann. Probab.* **14**, 1391–1398 (1986). Correction: *Ann. Probab.* **16**, 429–430 (1988)

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

[16] Vellaisamy, P.: Poisson approximation for (k_1, k_2)-events via the Stein-Chen method. *J. Appl. Probab.* **41**, 1081–1092 (2004)

[17] Wang, Y.H.: On the limit of the Markov binomial distribution. *J. Appl. Probab.* **18**, 937–942 (1981)