Ann Inst Stat Math (2015) 67:195–210
DOI 10.1007/s10463-013-0445-6

Compound Poisson approximation to weighted sums of symmetric discrete variables

A. Elijio · V. Čekanavičius

Received: 4 March 2013 / Revised: 25 October 2013 / Published online: 23 January 2014

Abstract The weighted sum $S = w_1 S_1 + w_2 S_2 + \cdots + w_N S_N$ is approximated by compound Poisson distribution. Here S_i are sums of symmetric independent identically distributed discrete random variables, and w_i denote weights. The estimates take into account the smoothing effect that sums S_i have on each other.

Keywords Concentration function · Compound Poisson distribution · Kolmogorov norm · Weighted random variables

1 Introduction

Let us consider the following complex sampling design: the entire population consists of different clusters, and the probability for each cluster to be selected into the sample is known. The design is usually referred to as cluster sampling and is often used in social surveys. The sum of sample elements then is equal to $S = w_1 S_1 + w_2 S_2 + \cdots + w_N S_N = w_1(X_{11} + X_{12} + \cdots + X_{1n_1}) + \cdots + w_N(X_{N1} + X_{N2} + \cdots X_{Nn_N})$. We assume further that all X_{ij}, $(i = 1, \ldots, N;\ \ j = 1, 2, \ldots, n_i)$ are independent and $X_{i1}, \ldots, X_{in_i}$ are identically distributed. Let F_i denote the distribution of $w_i X_{ij}$.

Many papers deal with the limiting behavior of weighted sums paying special attention to weights, see, for example, Liang and Baek (2006), Rosalsky and Sreehari (1998), Zhang (1997) and the references therein. In our paper, the emphasis is on random variables, which form a triangular array. We investigate approximation of S

A. Elijio · V. Čekanavičius (✉)
Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, Vilnius 03225, Lithuania
e-mail: vydas.cekanavicius@mif.vu.lt

A.Elijio
e-mail: aiste.elijio@gmail.com

by compound Poisson distribution taking into account smoothing effect that the whole sum has on its summands. This paper continues research of Čekanavičius and Elijio (2013).

To make our goals more explicit we need additional notation. Let $\mathcal{F}$ (resp. $\mathcal{S}, \mathcal{S}_+, \mathcal{M}$) denote the set of probability distributions (resp. symmetric probability distributions, distributions with nonnegative characteristic functions, finite signed measures) on $\mathbb{R}$. The Dirac measure concentrated at a is denoted by I_a, we use notation I for the Dirac measure concentrated at zero, that is $I \equiv I_0$. All products and powers of finite measures $W \in \mathcal{M}$ are defined in the convolution sense, and $W^0 = I$. The exponential of W is defined by $\exp\{W\} = \sum_{m=0}^{\infty} W^m/m!$. Note that Poisson (resp. compound Poisson) distribution can be expressed as $\exp\{\lambda(I_1 - I)\}$ (resp. $\exp\{\lambda(H - I)\}$, $H \in \mathcal{F}$). Note also that the distribution of S is $F_1^{n_1} F_2^{n_2} \cdots F_N^{n_N}$. We denote by $\widehat{W}(t)$ the Fourier–Stieltjes transform of $W \in \mathcal{M}$.

The Kolmogorov (uniform) norm $\| W \|_K$ and the total variation norm $\| W \|$ of $W \in \mathcal{M}$ are defined by

$$\| W \|_K = \sup_{x \in \mathbb{R}} |W\{(-\infty, x]\}|, \quad \| W \| = W^+\{\mathbb{R}\} + W^-\{\mathbb{R}\},$$

respectively. Here $W = W^+ - W^-$ is the Jordan–Hahn decomposition of W. If W is concentrated on $x_1, x_2, \ldots$, then $\| W \| = \sum_{k=1}^{\infty} |W\{x_k\}|$. Note that $\| W \|_K \leqslant \| W \|$, $\| WV \|_K \leqslant \| W \| \cdot \| V \|_K$ and for any distribution $\| F \| = 1$.

We explain motivating idea of this paper by considering $N = 2$. Let us assume that we want to approximate $F_1^{n_1} F_2^{n_2}$ by some compound Poisson distributions $G_1^{n_1} G_2^{n_2}$. Then by the triangle inequality and properties of the norms

$$\begin{aligned} \| F_1^{n_1} F_2^{n_2} - G_1^{n_1} G_2^{n_2} \|_K &\leqslant \| (F_1^{n_1} - G_1^{n_1}) F_2^{n_2} \|_K + \| G_1^{n_1} (F_2^{n_2} - G_2^{n_2}) \|_K \\ &\leqslant \| F_1^{n_1} - G_1^{n_1} \|_K + \| F_2^{n_2} - G_2^{n_2} \|_K. \end{aligned} \tag{1}$$

Such approach is reasonable only if both final estimates are of similar order. Otherwise, by neglecting $F_2^{n_2}$ in the first estimate, we can significantly worsen the overall accuracy of approximation. For example, if $n_1 = 1$, $n_2 = n$, then it is very likely for $\| F_1 - G_1 \|_K$ to be of constant order. Meanwhile, convolution with F_2^n can have strong smoothing effect. Our aim is to estimate this effect.

Technically the problem is quite challenging, since we can hardly use any of the standard approaches. In general, if we consider integer valued sums S_1 and S_2, the sums $w_1 S_1$ and $w_2 S_2$ are concentrated on different lattices. Consequently, a) no simple formula of inversion for $F_1^{n_1} F_2^{n_2}$ is available; b) it is impossible to apply Stein's method (moreover, we consider symmetric random variables).

We end this section by introducing remaining notation. Smoothing effect is estimated through Lévy's concentration function. For $F \in \mathcal{F}$, $h \geqslant 0$ Lévy's concentration function is defined by

$$Q(F, h) = \sup_x F\{[x, x + h]\}.$$

All absolute positive constants are denoted by C. Sometimes we supply C with indices. We use the notation θ for all quantities satisfying $|\theta| \leqslant 1$. For example, $u = \theta 3\mathrm{e}|t|$ means that $|u| \leqslant 3\mathrm{e}|t|$.

2 Known results

First of all, note partial cases of the first uniform Kolmogorov theorem

$$\sup_{F\in\mathcal{S}_+} \| F^n - \exp\{n(F-I)\} \|_K \leqslant C_1 n^{-1}, \quad \sup_{F\in\mathcal{S}} \| F^n - \exp\{n(F-I)\} \|_K \leqslant C_2 n^{-1/2}, \tag{2}$$

see Arak and Zaĭtsev (1988, pp. 116–117). For lattice distributions, there exists some analog of (2) in total variation norm. Let us consider symmetric three point random variable taking values ± 1 with probability $p < 1/4$ and zero with probability $1-2p$. Then

$$\| (I + p(I_1-I)+p(I_{-1}-I))^n - \exp\{np(I_1+I_{-1} - 2I)\} \| \leqslant C_3 \min(np^2, n^{-1}), \tag{3}$$

see, for example, Presman (1986, Eq. (11)). Estimates (2) and (3) show that symmetry of distributions significantly improves the accuracy of approximation. However, we cannot apply these estimates directly, since, in general, different $w_j S_j$ will never have the same symmetric component F. For the same reason, we cannot apply results for compound Poisson approximation of discrete symmetric distributions from Čekanavičius and Wang (2003), since in the mentioned paper all distributions have the same support.

If distributions have three finite absolute moments the Berry–Esseen theorem can be used:

$$\left\| \prod_{i=1}^{n} F_i - \Phi(\mu, \sigma^2) \right\|_K \leqslant \frac{C_4 \sum_{i=1}^{n} \beta_{3i}}{\left(\sum_{i=1}^{n} \sigma_i^2\right)^{3/2}}. \tag{4}$$

Here β_{3i} and σ_i^2 are the third absolute moment and variance of F_i, respectively. In many situations, (4) is sufficient for obtaining the required accuracy. Moreover, in some discrete cases of non-triangular arrays, second-order asymptotic expansion can be of the order $O(n^{-1})$, see Booth et al. (1994) However, this is not the case when random variables are close to zero, i.e. when Poisson approximation is more preferable.

In this paper, we consider F_i with sufficiently large probability mass concentrated at zero. Consequently, it allows decomposition $F_i = (1-p_i)I + p_i B_i$ for some $B_i \in \mathcal{F}$. Then

$$\left\| \prod_{i=1}^{n} \big((1-p_i)I+p_iB_i\big)-\exp\left\{\sum_{i=1}^{n} p_i(B_i-I)\right\} \right\|_K \leqslant C_5 \min\left\{\sum_{i=1}^{n} p_i^2, \max_i p_i\right\}. \tag{5}$$

The first estimate in (5) is known since the early 1930s and holds also for the total variation norm. Usually it is associated with the names of Khintchine, Doeblin and Le Cam. The estimate through maximal p_i is proved in Zaĭtsev (1984). In general, (5) cannot be improved. However, if part of B_i are equal, one can hope for a smoothing effect. Note also that, for nonsymmetric B_i, the estimate with smoothing effect has been obtained in Roos (2005).

3 Results

First we consider the case of discrete but not necessarily lattice symmetric distributions. Poisson type approximations for distributions with the same support have been considered in Čekanavičius (2003) and Čekanavičius and Wang (2003). We assume that all distributions have enough probability mass at zero: $F_j \in \mathcal{S}_+$. It is possible to apply (2) for each $F_j^{n_j}$ separately as outlined in (1). However, then the estimate will be of order $O(1/\min n_j)$, which can be quite rough. Further on we use notation $\sigma_i^2 = \int x^2 F_i\{\mathrm{d}x\}$ and assume that $s \geqslant 1$ is fixed.

Theorem 1 *Let $F_j \in \mathcal{S}_+$ be concentrated on a set $\{\pm w_j x_{j1}, \pm w_j x_{j2}, \dots, \pm w_j x_{js}\}$ $(j = 1, \dots, N)$, $n = \sum_{i=1}^N n_i$. Then, for any $h > 0$,*

$$\left\| \prod_{j=1}^{N} F_j^{n_j} - \exp\left\{\sum_{j=1}^{N} n_j\big(F_j - I\big)\right\} \right\|_K \leqslant C_6 Q\left(\exp\left\{\sum_{j=1}^{N} \frac{n_j}{2}\big(F_j - I\big)\right\}, h\right)$$
$$\times \left(\frac{1}{h}\sum_{j=1}^{N} \frac{\sigma_j}{\sqrt{n_j}} + C_7(N,s)\ln(n+1)\sum_{j=1}^{N}\frac{1}{n_j}\right) + e^{-n} \tag{6}$$

and

$$\left\| \prod_{j=1}^{N} F_j^{n_j} - \exp\left\{\sum_{j=1}^{N} n_j\big(F_j - I\big)\right\}\left(I - \frac{1}{2}\sum_{j=1}^{N} n_j\big(F_j - I\big)^2\right) \right\|_K$$
$$\leqslant C_8 Q\left(\exp\left\{\sum_{j=1}^{N} \frac{n_j}{2}\big(F_j - I\big)\right\}, h\right)$$
$$\times \left(\frac{1}{h}\sum_{j=1}^{N} \frac{\sigma_j}{n_j\sqrt{n_j}} + C_9(N,s)\ln(n+1)\sum_{j=1}^{N}\frac{1}{n_j^2}\right) + C_{10}e^{-n}. \tag{7}$$

For better appreciation of the smoothing effect, let us consider the following example. Let $n = n_1 + n_2$ and let all x_{ij} do not depend on n and let $F_i\{x_{ij}\}$ be uniformly bounded from zero and unity by some positive constants, $F_i\{x_{ij}\} \asymp C$. Then choosing $h = \min_{i=1,2;\, j\leqslant s} x_{ij}/2$ and applying (14) from Lemma 3 below, we get

$$\begin{aligned} &\| F_1^{n_1} F_2^{n_2} - \exp\{n_1(F_1 - I) + n_2(F_2 - I)\} \|_K \\ &\quad \leqslant \frac{C}{\sqrt{n}} \left(\frac{1}{\sqrt{n_1}} + \frac{1}{\sqrt{n_2}} + \ln(n+1) \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \right) + \mathrm{e}^{-n}. \end{aligned} \tag{8}$$

Note that direct application of the triangle inequality in a way similar to that of (1) combined with (2) provides the following estimate

$$\| F_1^{n_1} F_2^{n_2} - \exp\{n_1\big(F_1 - I\big) + n_2\big(F_2 - I\big)\} \|_K \leqslant C \left(\frac{1}{n_1} + \frac{1}{n_2} \right). \tag{9}$$

Consequently, the smoothing effect is realized through replacement of individual factors $n_j^{-1/2}$ by a common factor $n^{-1/2}$. This can significantly improve the accuracy. Indeed, let $n_1 = O(\sqrt{n_2})$. Then the order of accuracy in (8) is $O\big(n_2^{-3/4}\big)$. Meanwhile (9) gives $O(n_2^{-1/2})$. Moreover, application of the Berry-Esseen theorem (4) also gives the estimate of the order $O(n_2^{-1/2})$. Second order asymptotic expansion (7) improves the accuracy to $O(n_2^{-5/4})$.

Next we consider the case, when all $F_j \in \mathcal{S}$ are lattice distributions. However, due to possible differences in weights, F_j and F_k may have different lattice supports, when $j \neq k$. Unlike the previous case, supports can be infinite. We use standard decomposition $F_j = (1 - p_j)I + p_j B_j$. Let $\mu_{2j} = \int_{-\infty}^{\infty} x^2 B_j\{\mathrm{d}x\} = 2w_j^2 \sum_{k=1}^{\infty} k^2 B_j\{k\}$.

Theorem 2 *Let $B_j \in \mathcal{S}$ be concentrated on a set $\{\pm w_j, \pm 2w_j, \pm 3w_j, \ldots\}$, $\mu_{2j} < \infty$, and let $0 \leqslant p_j \leqslant \tilde{C} < 1$ $(j = 1, \ldots, N)$. Then, for all $h > 0$,*

$$\begin{aligned} &\left\| \prod_{j=1}^{N} ((1 - p_j)I + p_j B_j)^{n_j} - \exp\left\{ \sum_{j=1}^{N} n_j p_j \big(B_j - I\big) \right\} \right\|_K \leqslant C_{11} Q(M_1, h) \\ &\times \sum_{j=1}^{N} \left(n_j p_j^2 \sqrt{\mu_{2j}} \min\left\{ 1; \frac{1}{(n_j p_j)^{3/2}} \right\} \frac{1}{h} + \min\left\{ n_j p_j^2; \frac{\sqrt{\mu_{2j}}}{n_j} \right\} \right). \end{aligned} \tag{10}$$

Here $M_1 \in \mathcal{S}_+$ with the characteristic function $\widehat{M}_1(t) = \exp\Big\{\sum_{l=1}^{N} n_l(1 - p_l)$ $p_l\big(\widehat{B}_l(t) - 1\big)/2\Big\}$.

It is easy to check, that unlike (6), estimate (10) does not contain logarithmic factor. Moreover, F_j might not have any finite moment of order higher than 2.

Corollary 1 *Let in Theorem 2*

$$B_j = \frac{1}{2} I_{-w_j} + \frac{1}{2} I_{w_j}, \quad w_j \asymp C, \quad j = 1, \dots, N.$$

Then

$$\begin{aligned}
&\left\| \prod_{j=1}^{N} ((1-p_j)I + p_j B_j)^{n_j} - \exp\left\{ \sum_{j=1}^{N} n_j p_j \big(B_j - I\big) \right\} \right\|_K \\
&\leqslant C_{12} \min\left\{ 1, \left(\sum_{m=1}^{N} n_m p_m \right)^{-1/2} \right\} \sum_{j=1}^{N} \min\left(n_j p_j^2, \sqrt{\frac{p_j}{n_j}} \right). \qquad (11)
\end{aligned}$$

For the proof of (11) it suffices to take $h = \min_j w_j/2$ and apply (14). If $N = 1$, then (11) has the same order of accuracy as (3). Of course, the smoothing effect appears, when $N > 1$ and p_j, $(j = 1, \dots, N)$ tend to zero with different rates. Note also that Theorem 2 allows investigation of the impact of w_j on the accuracy of approximation. For example, we can assume that some of the weights converge to zero as the others are bounded by $\sqrt{n}$, etc.

Next we consider possible simplification of the structure of approximating compound Poisson distributions. Both previous theorems involve accompanying approximating distributions $\exp\{n_j(F_j - I)\}$. Such distributions can have complicated structures. From a practical point of view, the fewer the number of Poisson convolutions in Compound Poisson approximation, the more convenient it is for applications. Such reduction of number of convolutions, however, comes at a price of additional assumptions.

Though, in principle, we consider the same case as in Theorem 2, it is more convenient not to use its decomposition. Let $F_j \in \mathcal{S}$ be concentrated at $0,\ \pm w_j,\ \pm 2w_j,\ \pm 3w_j, \dots$, that is

$$F_j = p_{0j} I + \sum_{m=1}^{\infty} p_{mj} (I_{-mw_j} + I_{mw_j}), \quad p_{0j} + 2\sum_{m=1}^{\infty} p_{mj} = 1,\ \ j = 1, \dots, N. \qquad (12)$$

Inspired by Kruopis (1986) idea of left-handed and right-handed factorial moments we set

$$\nu_{mj} := \sum_{l=1}^{\infty} l(l-1)\dots(l-m+1) p_{lj}.$$

Poisson approximation is natural for random variables satisfying Franken's condition, see Franken (1964) and Kruopis (1986). We say that symmetric F_j satisfies Franken's condition if

$$\lambda_j := 2\nu_{1j} - (2\nu_{1j})^2 - 2\nu_{2j} = 2(\nu_{1j} - 2\nu_{1j}^2 - \nu_{2j}) > 0. \tag{13}$$

It is obvious that random variable satisfying (13) has most of its probability mass concentrated at zero.

Theorem 3 *Let F_j be defined by (12), $\int_{-\infty}^{\infty} x^4 F_j\{\mathrm{d}x\} < \infty$, and let (13) be satisfied for $j = 1, 2, \dots, N$. Then, for all $h > 0$, the following estimate holds*

$$\begin{aligned}
&\left\| \prod_{j=1}^{N} F_j^{n_j} - \exp\left\{ \sum_{j=1}^{N} n_j(\nu_{1j} + \nu_{2j})\big(I_{w_j} + I_{-w_j} - 2I\big) \right\} \right\|_K \\
&\quad \leqslant C_{13} Q(M_2, h) \sum_{j=1}^{N} n_j \big(\nu_{2j} + \nu_{3j} + \nu_{4j} + \nu_{1j}^2\big) \\
&\qquad \times \left(\frac{w_j}{h} \min\left\{1, \frac{1}{(\lambda_j n_j)^{3/2}}\right\} + \min\left\{1, \frac{1}{(\lambda_j n_j)^2}\left(1 + \frac{\nu_{1j} + \nu_{2j}}{\lambda_j}\right)\right\} \right).
\end{aligned}$$

Here $M_2 \in \mathcal{S}_+$ with $\widehat{M}_2(t) = \exp\left\{-\sum_{l=1}^{N} n_l \lambda_l \sin^2(t w_l/2)\right\}$.

We exemplify Theorem 3 assuming that $N = 2$ and w_1, w_2 are some absolute constants, $F_j = (1 - 4p_j)I + p_j(I_{-w_j} + I_{w_j}) + p_j(I_{-2w_j} + I_{2w_j})$, $p_j < 1/10$, $n_j p_j > 1$, $(j = 1, 2)$. Then $\lambda_j > p_j/5$, $\nu_{1j} = 3p_j$, $\nu_{2j} = 2p_j$, $\nu_{3j} = \nu_{4j} = 0$. Observe that

$$\begin{aligned}
M_2 = \frac{n_1\lambda_1 + n_2\lambda_2}{2} \bigg(&\frac{n_1\lambda_1}{2(n_1\lambda_1 + n_2\lambda_2)} I_{-w_1} + \frac{n_1\lambda_1}{2(n_1\lambda_1 + n_2\lambda_2)} I_{w_1} \\
&+ \frac{n_2\lambda_2}{2(n_1\lambda_1 + n_2\lambda_2)} I_{-w_2} + \frac{n_2\lambda_2}{2(n_1\lambda_1 + n_2\lambda_2)} I_{w_2} \bigg).
\end{aligned}$$

If $h > 0$ is a sufficiently small absolute constant, then from (14) below it follows that

$$Q(M_2, h) \leqslant \frac{C}{\sqrt{n_1\lambda_1 + n_2\lambda_2}} \leqslant \frac{C}{\sqrt{n_1 p_1 + n_2 p_2}}.$$

Therefore, choosing sufficiently small h, from Theorem 3 we get

$$\begin{aligned}
&\| F_1^{n_1} F_2^{n_2} - \exp\{5n_1 p_1(I_{-w_1} + I_{w_1} - 2I) + 5n_2 p_2(I_{-w_2} + I_{w_2} - 2I)\} \|_K \\
&\quad \leqslant \frac{C}{\sqrt{n_1 p_1 + n_2 p_2}} \left(\frac{1}{\sqrt{n_1 p_1}} + \frac{1}{\sqrt{n_2 p_2}} \right).
\end{aligned}$$

We see that the smoothing effect is similar to that of Theorem 1.

4 Auxiliary results

Let $\vec{u} = (u_1, \dots, u_{\bar{N}}) \in \mathbb{R}^{\bar{N}}$. Set

$$K_m(\vec{u}) = \left\{ \sum_{i=1}^{\bar{N}} j_i u_i : j_i \in \{-m, -m+1, \dots, m\}, i = 1, \dots, N \right\},$$

$$\delta(W, m, \vec{u}) = W^+\{\mathbb{R}\backslash K_m(\vec{u})\} + W^-\{\mathbb{R}\backslash K_m(\vec{u})\} = \int_{\mathbb{R}\backslash K_m(\vec{u})} 1\,|W\{\mathrm{d}x\}|.$$

Here W^+, W^- denote the positive and negative variation of W, respectively.

We will need the following lemmas.

Lemma 1 *Let* $W \in \mathcal{M}$, $W\{\mathbb{R}\} = 0$, $\bar{N}, m \in \mathbb{N}$, $\vec{u} \in \mathbb{R}^{\bar{N}+1}$, $h > 0$ *and* $U \in \mathcal{F}_+$. *Then*

$$\| W \|_K \leqslant C \int_{|t| \leqslant 1/h} \left| \frac{\widehat{W}(t)}{t} \right| \mathrm{d}t + C(N) \ln(m+1) \sup_{t \in \mathbb{R}} \frac{|\widehat{W}(t)|}{\widehat{U}(t)} Q(U, h) + \delta(W, m, \vec{u}).$$

Lemma 2 *Let* $V_1, V_2 \in \mathcal{M}$ *be such that* $\|V_1\| \leqslant b_1$, $\|V_2\| \leqslant b_2$ *and, for some* $s, \bar{N}, m \in \mathbb{N}$ *and* $\vec{u} \in \mathbb{R}^{\bar{N}+1}$, *let supp* $V_1 \subset K_s(\vec{u})$ *and supp* $V_2 \subset K_m(\vec{u})$. *Then, for all* $y \in \mathbb{N}$,

$$\delta(V_1 \exp\{V_2\}, s + my, \vec{u}) \leqslant b_1 \exp\{3b_2 - y\}, \quad \delta(\exp\{V_2\}, my, \vec{u}) \leqslant \exp\{3b_2 - y\}.$$

Lemmas 1 and 2 have been proved in Čekanavičius and Wang (2003).

Lemma 3 *Let* $F, G \in \mathcal{F}$, $h > 0$ *and* $a > 0$. *Then*

$$Q(\exp\{a(F - I)\}, h) \leqslant \frac{C}{\sqrt{aF\{|x| > h\}}}. \tag{14}$$

If, in addition, $\widehat{F}(t) \geqslant 0$, *then*

$$h \int_{|t| \leqslant 1/h} |\widehat{F}(t)|\,\mathrm{d}t \leqslant C Q(F, h). \tag{15}$$

Lemma 3 contains well-known properties of Levy's concentration function, see, for example, Arak and Zaĭtsev 1988, Chapter 2.

One of the main tools used in this paper is the following modified Le Cam's inequality from Čekanavičius and Elijio (2013) (see also Le Cam 1965; Čekanavičius and Roos 2006a).

Lemma 4 *Let* $h > 0$, $W \in \mathcal{M}$, $W\{\mathbb{R}\} = 0$, $H \in \mathcal{F}$, $M \in \mathcal{S}_+$, *and* $|\widehat{H}(t)| \leqslant C\widehat{M}(t)$, *for* $|t| \leqslant 1/h$. *Then*

$$\| WH \|_K \leqslant C \left(\sup_{|t| \leqslant 1/h} \frac{|\widehat{W}(t)|}{|t|} \cdot \frac{1}{h} + \|W\| \right) Q(M, h).$$

Next two Lemmas have been proved in Čekanavičius and Roos (2006a, b).

Lemma 5 *Let $j \in \{1, 2, \dots\}$, $n \in \mathbb{N}$, and $p = 1-q \in (0, 1)$. If $F \in \mathcal{S}$ is concentrated on the set $\{\pm 1, \pm 2, \dots\}$ and has finite variance σ^2, then*

$$\| (F-I)^j (I + p(F-I))^n \| \leqslant 6.73 \sqrt{\sigma}\, \frac{j}{q^{1/4}} \left(\frac{j}{\mathrm{e}\, npq} \right)^j .$$

Lemma 6 *Let $F \in \mathcal{S}$ be concentrated on $\{\pm 1, \pm 2, \dots\}$ and has a finite variance σ^2. Let $a \in (0, \infty)$, $k \in \mathbb{N}$. Then*

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

Lemma 7 *Let $\widehat{R}(t) = \sum_{k=-\infty}^{\infty} R_k \mathrm{e}^{itk}$, $\sum_{k=-\infty}^{\infty} |k||R_k| < \infty$. Then, for all $\gamma > 0$ and $\upsilon \in R$,*

$$\left(\sum_{k=-\infty}^{\infty} |R_k| \right)^2 \leqslant \left(\frac{1}{2} + \frac{1}{2\pi\gamma} \right) \int_{-\pi}^{\pi} \left(\gamma |\widehat{R}(t)|^2 + \frac{1}{\gamma} \left| \left(\widehat{R}(t) \mathrm{e}^{-it\upsilon} \right)' \right|^2 \right) \mathrm{d}t.$$

Lemma 7 has been proved in Presman (1986).

5 Proofs

For the sake of brevity, we omit the dependence of Fourier transforms on t whenever it does not lead to ambiguity. For example, we write $\widehat{F}_i$ instead of $\widehat{F}_i(t)$. Let $G_i = \exp\{F_i - I\}$ and let Θ denote all measures, satisfying $\| \Theta \| \leqslant 1$. We also constantly use two simple estimates

$$x^k \mathrm{e}^{-ax^2} \leqslant C(k) a^{-k/2}, \qquad \left| \mathrm{e}^b - \sum_{j=0}^{k} \frac{b^j}{j!} \right| \leqslant \frac{|b|^{k+1}}{(k+1)!}, \tag{16}$$

which hold, for any $x > 0$, $a > 0$, $b \leqslant 0$ and $k \in \mathbb{N}$.

Proof of Theorem 1 Proofs of both estimates, (6) and (7), are based on Lemmas 1 and 2. First we use Lemma 1. In this case, set $W = \prod_{i=1}^{N} F_i^{n_i} - \exp\left\{ \sum_{i=1}^{N} n_i \left(F_i - I \right) \right\}$ and $U = \exp\left\{ \sum_{j=1}^{N} n_j \left(F_j - I \right)/2 \right\}$.

It is obvious, that $\widehat{F}_i \leqslant \exp\{\widehat{F}_i - 1\}$, since $0 \leqslant \widehat{F}_i \leqslant 1$. Moreover, $\exp\{1 - \widehat{F}_i\} \leqslant \mathrm{e}$. Therefore, taking into account (16), it is not difficult to prove that

$$|\widehat{F}_i^{n_i} - \widehat{G}_i^{n_i}| \leqslant n_i |\widehat{F}_i - \widehat{G}_i| \widehat{G}_i^{n_i - 1} \leqslant C \widehat{G}_i^{n_i} n_i |\widehat{F}_i - 1|^2$$

and

$$|\widehat{W}| \leqslant \left|\prod_{i=1}^{N} \widehat{F}_i^{n_i} - \prod_{i=1}^{N} \widehat{G}_i^{n_i}\right| \leqslant C\widehat{U}^2 \sum_{i=1}^{N} n_i |\widehat{F}_i - \widehat{G}_i| \leqslant C\widehat{U}^2 \sum_{i=1}^{N} n_i |\widehat{F}_i - 1|^2.$$

From that we get two important inequalities. By the first estimate in (16),

$$|\widehat{W}| \leqslant C\widehat{U} \sum_{i=1}^{N} n_i |\widehat{F}_i - 1|^2 \exp\{(n_i/2)(\widehat{F}_i - 1)\} \leqslant C\widehat{U}(t) \sum_{i=1}^{N} \frac{1}{n_i}.$$

Similarly

$$|\widehat{W}| \leqslant C\widehat{U}^2 \sum_{i=1}^{N} n_i |\widehat{F}_i - 1|^{3/2} \sigma_i |t| \leqslant C\widehat{U} \sum_{i=1}^{N} \frac{\sigma_i |t|}{\sqrt{n_i}}, \qquad \frac{|\widehat{W}|}{|t|} \leqslant C\widehat{U} \sum_{i=1}^{N} \frac{\sigma_i}{\sqrt{n_i}}.$$

Substituting these estimates into Lemma 1 and applying (15), we obtain

$$|W| \leqslant CQ(U, h) \left(\frac{1}{h} \sum_{j=1}^{N} \frac{\sigma_j}{\sqrt{n_j}} + C(N) \ln(m+1) \sum_{j=1}^{N} \frac{1}{n_j} \right) + \delta(W, m, \vec{u}).$$

We still need to estimate $\delta(W, m, \vec{u})$ for suitably chosen m and $\vec{u}$. For that purpose we will use Lemma 2. In this particular case, let $V_1 = \prod_{j=1}^{N} F_j^{n_j}$ and $V_2 = \sum_{j=1}^{N} n_j \big(F_j - I\big)$.

Distribution F_i is concentrated on $\{\pm w_i x_{i1}, \pm w_i x_{i2}, \ldots, \pm w_i x_{is}\}$. By choosing $\vec{u}_i = (0, w_i \vec{x}_i) = (0, w_i x_{i1}, \ldots, w_i x_{is})$, we get that $supp\, F_i \subset K_1(\vec{u}_i)$ and $supp\, F_i^{n_i} \subset K_{n_i}(\vec{u}_i)$. Here $supp F_i$ denotes the support of F_i. Let $\vec{u} := (0, w_1\vec{x}_1, w_2\vec{x}_2, \ldots, w_N\vec{x}_N) \in \mathbb{R}^{Ns+1}$. Obviously, $supp\, V_1 = supp\, F_1^{n_1} \cdots F_N^{n_N} \subset K_{n_1+\ldots+n_N}(\vec{u}) = K_n(\vec{u})$.

From the definition of V_2, we easily get $\|V_2\| \leqslant 2(n_1 + \ldots + n_N) = 2n$, and $supp\, V_2 \subset K_1(\vec{u})$. Therefore, from Lemma 2 it follows that $\delta(\mathrm{e}^{V_2}, y, \vec{u}) \leqslant \mathrm{e}^{6n-y}$. Let $y = 7n$. Then $\delta(e^{V_2}, 7n, \vec{u}) \leqslant e^{-n}$. From the definition of δ, we also have that $\delta(V_1, n, \vec{u}) = 0$. Therefore, it follows that $\delta(V_1 + \mathrm{e}^{V_2}, 7n, \vec{u}) \leqslant \mathrm{e}^{-n}$. Since $\delta(W, 7n, \vec{u}) = \delta(V_1 - \mathrm{e}^{V_2}, 7n, \vec{u}) \leqslant \delta(V_1 + \mathrm{e}^{V_2}, 7n, \vec{u}) \leqslant \mathrm{e}^{-n}$, we now easily get (6) by choosing $m = 7n$.

The estimate (7) is proved similarly. Taking into account (3.16)–(3.18) from Čekanavičius and Wang (2003), we obtain

$$|\widehat{W}| := \left| \prod_{j=1}^{N} \widehat{F}_j^{n_j} - \exp\left\{ \sum_{j=1}^{N} n_j \big(\widehat{F}_j - 1\big) \right\} \left(1 - \frac{1}{2} \sum_{k=1}^{N} n_k \big(\widehat{F}_k - 1\big)^2 \right) \right|$$

$$\leqslant C\widehat{U}^2 \left(\sum_{j=1}^{N} n_j |\widehat{F}_j - 1|^3 + \left(\sum_{j=1}^{N} n_j (\widehat{F}_j - 1)^2 \right)^2 \right) \leqslant C(N)\widehat{U}(t) \sum_{j=1}^{N} \frac{1}{n_j^2}$$

and

$$|\widehat{W}| \leqslant C\widehat{U}\sum_{j=1}^{N}\frac{\sigma_i|t|}{n_j\sqrt{n_j}}.$$

It remains to estimate $\delta(W, m, \vec{u})$. In this case, we choose the same $\vec{u}$ as before and take $V_3 = \big(I - \frac{1}{2}\sum_{j=1}^{N} n_j(F_j - I)^2\big)$. We have

$$\| V_3 \| \leqslant 1 + \frac{1}{2}\sum_{j=1}^{N} n_j \cdot 2^2 = 1 + 2n, \quad \text{supp } V_3 \subset K_1(\vec{u}).$$

Taking $y = 8n$, $m = 1$, from Lemma 2 we obtain

$$\delta(V_3 \mathrm{e}^{V_2}, 9n, \vec{u}) \leqslant (2n+1)e^{6n-y} \leqslant C\mathrm{e}^{-n}$$

and $\delta(V_1 - V_3\mathrm{e}^{V_2}, 9n, \vec{u}) \leqslant \delta(V_1, 9n, \vec{u}) + \delta(V_3\mathrm{e}^{V_2}, 9n, \vec{u})$. Applying Lemmas 1 and 2 we complete the proof of (7). □

Proof of Theorem 2 We give the proof for the case $N = 2$ only. For the general case, the argument is the same. Let $F_j = (1 - p_j)I + p_j B_j$. Let $\lfloor x \rfloor$ denote integer part of x $(0 \leqslant x - \lfloor x \rfloor < 1)$. Let $W_j = (F_j - I)^2 F_j^{\lfloor (k-1)/2 \rfloor} G_j^{\lfloor (n_j - k)/2 \rfloor}$ and let

$$H_1 = F_1^{\lfloor (k-1)/2 \rfloor} G_1^{\lfloor (n_1 - k)/2 \rfloor} F_2^{n_2}, \qquad H_2 = F_2^{\lfloor (k-1)/2 \rfloor} G_2^{\lfloor (n_2 - k)/2 \rfloor} G_1^{n_1}. \tag{17}$$

Taking into account the definition of G_j we obtain

$$\begin{aligned} F_j - G_j &= I + (F_j - I) - \sum_{k=0}^{\infty}\frac{(F_j - I)^k}{k!} = -(F_j - I)^2 \sum_{k=2}^{\infty}\frac{(F_j - I)^{k-2}}{k!} \\ &= (F_j - I)^2 \Theta C. \end{aligned}$$

Then

$$\begin{aligned} &\| F_1^{n_1} F_2^{n_2} - G_1^{n_1} G_2^{n_2} \|_K \\ &\leqslant \| (F_1^{n_1} - G_1^{n_1}) F_2^{n_2} \|_K + \| (F_2^{n_2} - G_2^{n_2}) G_1^{n_1} \|_K \\ &\leqslant \sum_{k=1}^{n_1} \left\| (F_1 - G_1) F_1^{k-1} G_1^{n_1 - k} F_2^{n_2} \right\|_K + \sum_{k=1}^{n_2} \left\| (F_2 - G_2) F_2^{k-1} G_2^{n_2 - k} G_1^{n_1} \right\|_K \\ &\leqslant C \| W_1 H_1 \|_K + C \| W_2 H_2 \|_K . \end{aligned} \tag{18}$$

We apply Lemma 4 to $W_j H_j$. First note that

$$\widehat{F}_j^2 = (1 - p_j)^2 + 2p_j(1 - p_j)\widehat{B}_j + p_j^2 \widehat{B}_j^2 \leqslant 1 + 2p_j(1 - p_j)(\widehat{B}_j - 1).$$

Therefore,

$$|\widehat{F}_j|, |\widehat{G}_j| \leqslant \exp\{p_j(1-p_j)(\widehat{B}_j - 1)\}. \tag{19}$$

Note that

$$\lfloor (k-1)/2 \rfloor + \lfloor (n_j - k)/2 \rfloor \geqslant \frac{k-1}{2} - 1 + \frac{n_j - k}{2} - 1 \geqslant \frac{n_j}{2} - \frac{5}{2}.$$

Consequently,

$$\begin{aligned} |\widehat{H}_1| &\leqslant \exp\left\{ n_2 p_2 (1-p_2)(\widehat{B}_2 - 1) + \frac{n_1}{2} p_1 (1-p_1)(\widehat{B}_1 - 1) + \theta \frac{5}{2} p_1 (1-p_1) 2 \right\} \\ &\leqslant C \widehat{M}_1(t). \end{aligned} \tag{20}$$

Similarly $|\widehat{H}_2| \leqslant C\widehat{M}_2(t)$.

We have

$$|\widehat{F}_j^{\lfloor (k-1)/2 \rfloor} \widehat{G}_j^{\lfloor (n_j - k)/2 \rfloor}| \leqslant C \exp\{C n_j p_j (\widehat{B}_j - 1)\}. \tag{21}$$

Indeed, if $n_j \geqslant 20$ the estimate follows from (19). If $n_j < 20$, then we make use of the fact, that any characteristic function is bounded by unity and, therefore,

$$\begin{aligned} |\widehat{F}_j^{\lfloor (k-1)/2 \rfloor} \widehat{G}_j^{\lfloor (n_j - k)/2 \rfloor}| &\leqslant 1 \leqslant \exp\{n_j p_j (\widehat{B}_j - 1)\} \exp\{20 p_j (|\widehat{B}_j| + 1)\} \\ &\leqslant \mathrm{e}^{40} \exp\{n_j p_j (\widehat{B}_j - 1)\}. \end{aligned}$$

Standard expansion of characteristic function results in inequality

$$|\widehat{F}_j - 1|^2 = p_j^2 |\widehat{B}_j - 1|^2 \leqslant C p_j^2 \sqrt{\mu_{2j}} |t| |\widehat{B}_j - 1|^{3/2}.$$

Combining the last estimate with (21) and (16) we obtain

$$\begin{aligned} \frac{|\widehat{W}_j|}{|t|} &\leqslant C p_j^2 \sqrt{\mu_{2j}} (1 - \widehat{B}_j)^{3/2} \exp\{C n_j p_j (\widehat{B}_j - 1)\} \\ &\leqslant C p_j^2 \sqrt{\mu_{2j}} \min(1, (n_j p_j)^{-3/2}). \end{aligned} \tag{22}$$

It remains to estimate $\| W_j \|$. It is obvious that $\| W \| \leqslant C p_j^2$. We recall that total variation norm of any distribution equals unity. Therefore, if $n_j \geqslant 20$ and $k > n_j/2$, then by Lemma 5

$$\| W_j \| \leqslant p_j^2 \| (B_j - I)^2 (I + p_j (B_j - I))^{\lfloor (k-1)/2 \rfloor} \| \leqslant C \frac{\sqrt{\mu_{2j}}}{n_j^2}.$$

Similar estimate is obtained via Lemma 6, if $n_j \geqslant 20$ and $k \leqslant n_2/2$. Now, let $n_j < 20$. Then

$$\begin{aligned}\| W_j \| &\leqslant Cp_j^2 \| (B_j - I)^2 \| = Cp_j^2 \| (B_j - I)^2 \exp\{n_j p_j (B_j - I)\} \exp\{n_j p_j (I - B_j)\} \| \\ &\leqslant Cp_j^2 \| (B_j - I)^2 \exp\{n_j p_j (B_j - I)\} \| \exp\{n_j p_j \| B_j - I \|\} \\ &\leqslant Cp_j^2 \| (B_j - I)^2 \exp\{n_j p_j (B_j - I)\} \| \leqslant C \frac{\sqrt{\mu_{2j}}}{n_j^2}. \end{aligned} \tag{23}$$

Combining last estimates with (22), (20), Lemma 4 and substituting the resulting estimate into (18) we complete the proof of Theorem 2. □

Proof of Theorem 3 We apply similar argument as used for the proof of Theorem 2. We use (17) and (18) with F_j defined by (12) and G_j with $\widehat{G}_j(t) = \exp\{(\nu_{1j} + \nu_{2j})(e^{it_j} - 1 + e^{-it_j} - 1)\}$. Hence and further $t_j \equiv t w_j$, $r_j = \nu_{2j} + \nu_{3j} + \nu_{4j} + \nu_{1j}^2$. Note that $\nu_{1j}, \nu_{2j} \leqslant 1$, since $\lambda_j > 0$. Applying Lemma 2 from Šiaulys and Čekanavičius (1988), we obtain

$$|\widehat{F}_j| \leqslant \exp\left\{-2\lambda_j \sin^2 \frac{t_j}{2}\right\}, \quad |\widehat{G}_j| \leqslant \exp\left\{-2\lambda_j \sin^2 \frac{t_j}{2}\right\}. \tag{24}$$

Taking into account the last estimates and arguing just like in the proof of (20) and (21) we get

$$|\widehat{H}_j| \leqslant \widehat{M}_2(t), \quad |\widehat{F}_j^{\lfloor (k-1)/2 \rfloor} \widehat{G}_j^{\lfloor (n_j - k)/2 \rfloor}| \leqslant C \exp\{-C n_j \lambda_j \sin^2(t_j/2)\}. \tag{25}$$

Next note that, for any $t \in \mathbb{R}$,

$$\begin{aligned}(\mathrm{e}^{\mathrm{i}t} - 1) + (\mathrm{e}^{-\mathrm{i}t} - 1) &= \mathrm{e}^{-\mathrm{i}t}(\mathrm{e}^{\mathrm{i}t} - 1)^2, \quad (\mathrm{e}^{\mathrm{i}t} - 1)^3 + (\mathrm{e}^{-\mathrm{i}t} - 1)^3 \\ &= (\mathrm{e}^{\mathrm{i}t} - 1)^4 \mathrm{e}^{-3\mathrm{i}t}(\mathrm{e}^{2\mathrm{i}t} + \mathrm{e}^{\mathrm{i}t} + 1), \; (\mathrm{e}^{\mathrm{i}t} - 1)^2 + (\mathrm{e}^{-\mathrm{i}t} - 1)^2 \\ &= 2(\mathrm{e}^{\mathrm{i}t} + \mathrm{e}^{-\mathrm{i}t} - 2) + \mathrm{e}^{-2\mathrm{i}t}(\mathrm{e}^{\mathrm{i}t} - 1)^4.\end{aligned}$$

Therefore, applying standard expansion in factorial moments, we obtain

$$\begin{aligned}\widehat{F}_j &= 1 + \nu_{1j}(\mathrm{e}^{\mathrm{i}t_j} + \mathrm{e}^{-\mathrm{i}t_j} - 2) + \frac{\nu_{2j}}{2}((\mathrm{e}^{\mathrm{i}t_j} - 1)^2 + (\mathrm{e}^{-\mathrm{i}t_j} - 1)^2) \\ &\quad + \frac{\nu_{3j}}{6}((\mathrm{e}^{\mathrm{i}t_j} - 1)^3 + (\mathrm{e}^{-\mathrm{i}t_j} - 1)^3) + \theta C \nu_{4j} |\mathrm{e}^{\mathrm{i}t_j} - 1|^4 \\ &= 1 + (\nu_{1j} + \nu_{2j})(\mathrm{e}^{\mathrm{i}t_j} + \mathrm{e}^{-\mathrm{i}t_j} - 2) + \theta C(\nu_{2j} + \nu_{3j} + \nu_{4j})|\mathrm{e}^{\mathrm{i}t_j} - 1|^4. \end{aligned} \tag{26}$$

Similarly,

$$\widehat{G}_j = 1 + (\nu_{1j} + \nu_{2j})(\mathrm{e}^{\mathrm{i}t_j} + \mathrm{e}^{-\mathrm{i}t_j} - 2) + \theta C(\nu_{2j} + \nu_{1j})^2 |\mathrm{e}^{\mathrm{i}t_j} - 1|^4. \tag{27}$$

Therefore,

$$|\widehat{F}_j - \widehat{G}_j| \leqslant C\theta|e^{it_j} - 1|^4 r_j$$

and

$$\frac{|W_j|}{|t|} \leqslant Cr_j w_j \min(1, (n_j\lambda_j)^{-3/2}). \tag{28}$$

It remains to estimate $\| W_j \|$. The total variation norm is invariant to scale change. Therefore, without loss of generality, we assume that W_j is concentrated at nonnegative integers, and $t_j = t$. Expanding F_j in powers of $(I_1 - I)$ (see, for example, Lemma 2.3 from Čekanavičius 1998) we can write analogs of (26) and (27):

$$F_j = I + (\nu_{1j} + \nu_{2j})(I_1 + I_{-1} - 2I) + (I_1 - I)^4\Theta C(\nu_{2j} + \nu_{3j} + \nu_{4j}),$$
$$G_j = I + (\nu_{1j} + \nu_{2j})(I_1 + I_{-1} - 2I) + (I_1 - I)^4\Theta C(\nu_{2j} + \nu_{1j})^2$$

and

$$F_j - G_j = (I_1 - I)^4\Theta C r_j.$$

It is obvious, that $\| W_j \| \leqslant Cr_j$. Therefore, it remains to prove estimate containing $C(n_j a_j)^{-2}$. Let us assume that $k < n_j/2$ and $n_j \geqslant 20$. Then

$$\begin{aligned}\| W_j \| &\leqslant Cr_j \| (I_1 - I)^4 \exp\{Cn_j\lambda_j(I_1 + I_{-1} - 2I)\} \| \\ &\leqslant Cr_j \| (I_1 - I)^4 \exp\{Cn_j\lambda_j(I_1 - I)\} \| \\ &\leqslant Cr_j \| (I_1 - I) \exp\{Cn_j\lambda_j(I_1 - I)/4\} \|^4 \leqslant Cr_j(n_j\lambda_j)^{-2}.\end{aligned}$$

For the last estimate we used well-known inequality (see, for example, Ibragimov and Presman 1973, Eq. (28); or Čekanavičius and Roos 2006b, Lemma 4.1). Now let us assume that $k \geqslant n_j/2$ and $n_j \geqslant 20$. For the sake of convenience we use abbreviation $a_j = \lfloor (k-1)/2 \rfloor$. Note that $a_j \geqslant n_j/6$. Let $n_j a_j \geqslant 1$. By the properties of total variation norm

$$\| W_j \| \leqslant Cr_j \| (I_1 - I)^4 F_j^{a_j} \|.$$

We will use Lemma 7 with $\gamma = \sqrt{n_j a_j}$, $\upsilon = 0$ and

$$\widehat{R}(t) = \widehat{F}_j^{a_j}(\mathrm{e}^{it} - 1)^4.$$

Taking into account (24) we obtain

$$|\widehat{R}(t)| \leqslant C\exp\{-2a_j\lambda_j \sin^2(t/2)\}\sin^4(t/2) \leqslant C(n_j\lambda_j)^{-2}\exp\{-a_j \sin^2(t/2)\}.$$

From Šiaulys and Čekanavičius (1988, Lemma 4) it follows that

$$|\widehat{F}_j'| \leqslant C(\nu_{1j}+\nu_{2j})|\mathrm{e}^{\mathrm{i}t}-1|.$$

Therefore,

$$\begin{aligned}|R'(t)| &\leqslant 4|\mathrm{e}^{\mathrm{i}t}-1|^3|\widehat{F}_j^{a_j}|+|\mathrm{e}^{\mathrm{i}t}-1|^4 a_j|\widehat{F}_j|^{a_j-1}|\widehat{F}_j'|\\ &\leqslant C\exp\{-2a_j\lambda_j\sin^2(t/2)\}\left(|\sin(t/2)|^3+a_j(\nu_{1j}+\nu_{2j})|\sin(t/2)|^5\right)\\ &\leqslant C\exp\{-a_j\lambda_j\sin^2(t/2)\}(n_j\lambda_j)^{-3/2}\left(1+\frac{\nu_{1j}+\nu_{2j}}{\lambda_j}\right).\end{aligned}$$

It remains to substitute the last two estimates in Lemma 7 to obtain estimate

$$\| F_j^{a_j}(I_1-I)^4 \| \leqslant C(n_j\lambda_j)^{-2}\left(1+\frac{\nu_{1j}+\nu_{2j}}{\lambda_j}\right).$$

For the case $n_j < 20$ argument goes exactly as in (23). Thus, we proved that

$$\| W_j \| \leqslant Cr_j \min\left(1, (n_j\lambda_j)^{-2}\left(1+\frac{\nu_{1j}+\nu_{2j}}{\lambda_j}\right)\right). \tag{29}$$

Consequently substituting estimates (29), (28), (25) into Lemma 4 and the resulting estimate in (18) we complete the proof of Theorem. □

Acknowledgments The authors wish to thank the referee for constructive suggestions which improved the paper.

References

Arak, T. V., Zaĭtsev, A. Yu. (1988). Uniform limit theorems for sums of independent random variables. *Proceedings of the Steklov Institute of Mathematics*, *188*(1), 1–222.

Booth, J. G., Hall, P., Wood, A. T. A. (1994). On the validity of Edgeworth and saddlepoint approximations. *Journal of Multivariate Analysis*, *51*, 121–138.

Čekanavičius, V. (1998). Estimates in total variation for convolutions of compound distributions. *Journal of London Mathematical Society*, *58*(2), 748–760.

Čekanavičius, V. (2003). Infinitely divisible approximations for discrete nonlattice variables. *Advances in Applied Probability*, *35*, 982–1006.

Čekanavičius, V., Elijio, A. (2013). Smoothing effect of Compound Poisson approximation to distribution of weighted sums. *Lithuanian Mathematical Journal* (to appear, preprint version is at arXiv:1303.0159).

Čekanavičius, V., Roos, B. (2006a). An expansion in the exponent for compound binomial approximations. *Lithuanian Mathematical Journal*, *46*, 54–91.

Čekanavičius, V., Roos, B. (2006b). Compound binomial approximations. *Annals of the Institute of Statistical Mathematics*, *58*(1), 187–210.

Čekanavičius, V., Wang, Y. H. (2003). Compound Poisson approximations for sums of discrete nonlattice variables. *Advances in Applied Probability*, *35*, 228–250.

Franken, P. (1964). Approximation des Verteilungen von Summen unabhängiger nichtnegativer ganzzähliger Zufallsgrößen durch Poissonsche Verteilungen. *Mathematische Nachrichten*, *27*, 303–340.

Ibragimov, I. A., Presman, É. L. (1973). On the rate of approach of the distributions of sums of independent random variables to accompanying distributions. *Theory of Probability and Its Applications*, *18*, 713–727.

Kruopis, J. (1986). Approximations for distributions of sums of lattice random variables I. *Lithuanian Mathematical Journal*, *26*, 234–244.

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

Liang, H.-Y., Baek, J.-I. (2006). Weighted sums of negatively associated random variables. *Australian and New Zealand Journal of Statistics*, *48*(1), 21–31.

Presman, E. L. (1986). Approximation in variation of the distribution of a sum of independent Bernoulli variables with a Poisson law. *Theory of Probability and Its Applications*, *30*(2), 417–422.

Roos, B. (2005). On Hipp's compound Poisson approximations via concentration functions. *Bernoulli*, *11*(3), 533–557.

Rosalsky, A., Sreehari, M. (1998). On the limiting behavior of randomly weighted partial sums. *Statistics and Probability Letters*, *40*, 403–410.

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

Zaĭtsev, A. Yu. (1984). On the accuracy of approximation of distributions of sums of independent random variables-which are nonzero with a small probability-by means of accompanying laws. *Theory of Probability and Its Applications*, *28*(4), 657–669.

Zhang, Li-Xin. (1997). Strong approximation theorems for geometrically weighted random series and their applications. *Annals of Probability*, *25*(4), 1621–1635.