Smoothing effect of Compound Poisson approximation to distribution of weighted sums

V. Čekanavičius and A. Elijio

Department of Mathematics and Informatics, Vilnius University,

Naugarduko 24, Vilnius 03225, Lithuania.

E-mail: vydas.cekanavicius@mif.vu.lt and aiste.elijio@gmail.com

Abstract

The accuracy of compound Poisson approximation to the sum $S = w_1S_1+w_2S_2+\cdots+w_NS_N$ is estimated. Here S_i are sums of independent or weakly dependent random variables, and w_i denote weights. The overall smoothing effect of S on w_iS_i is estimated by Lévy concentration function.

Key words: *characteristic function, concentration function, compound Poisson distribution, Kolmogorov norm, weighted random variables.*

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

1 Introduction

Let us consider typical cluster sampling design: the entire population consists of different clusters, and the probability for each cluster to be selected into the sample is known. The sum of sample elements then is equal to $S = w_1S_1 + w_2S_2 + \cdots + w_NS_N = w_1(X_{11} + X_{12} + \cdots + X_{1n_1}) + \cdots + w_N(X_{N1} + X_{N2} + \cdots X_{Nn_N})$. Here w_i denote weights, which are inversely proportional to probabilities to be selected into sample.

We explain motivating idea of this paper by considering simple example, when $N = 2$ and $w_1 = w_2 = 1$. We want to estimate $d(S_1 + S_2, Z_1 + Z_2)$, where $d(\cdot,\cdot)$ denotes some probabilistic metric. The majority of metrics allows the following simplification

$$d(S_1 + S_2, Z_1 + Z_2) \leqslant d(S_1 + S_2, Z_1 + S_2) + d(Z_1 + S_2, Z_1 + Z_2) \leqslant d(S_1, Z_1) + d(S_2, Z_2). \qquad (1)$$

Such approach is reasonable only if both final estimates are of similar order. Otherwise, by neglecting S_2, we can significantly worsen the overall estimate of the accuracy of approximation. For example, let S_1 have just few summands and $d(S_1, Z_1) = O(1)$. Let S_2 have a large number of summands. Then, neither $S_1 + S_2$ nor $Z_1 + S_2$ differ much from S_2 and, it is natural to expect $d(S_1 + S_2, Z_1 + S_2)$ to be small. If this is the case, we say that S_2 has smoothing effect on S_1. Our aim is investigation of such smoothing effects.

Weighting can radically change the structural properties of S. For example, even if all S_i are lattice, the sum S is not necessarily lattice random variable. Therefore, the standard approaches (Tsaregradski's inequality, Stein's method) are inapplicable.

We introduce necessary notation. Let $\mathcal{F}$ (resp. $\mathcal{M}$) denote the set of probability distributions (resp. finite signed measures) on $\mathbb{R}$. The Dirac measure concentrated at a is denoted by I_a, $I = I_0$. All products and powers of finite signed measures $W \in \mathcal{M}$ are defined in the convolution sense, and $W^0 = I$. The exponential of W is the finite signed measure defined by $\exp\{W\} = \sum_{m=0}^{\infty} W^m/m!$. 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. Note that $\| W \|_K \leqslant \| W \|$, $\| WV \|_K \leqslant \| W \| \cdot \| V \|_K$. If $F \in \mathcal{F}$, then $\| F \|_K = \| F \| = 1$. 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 the same symbol C. Sometimes we supply C with indices. We also assume usual convention $\sum_{j=a}^{b} = 0$ and $\prod_{j=a}^{b} = 1$, if $b < a$.

2 Known results

As a rule, the limiting behavior of weighted sums is investigated with the emphasis on weights, for example, see [12], [17], [20] and the references therein. In our paper, emphasis is on the structure of random variables.

Let us assume that all distributions have finite thee absolute moments. Then the Berry-Esseen theorem can be used:

$$\Big\| \prod_{i=1}^{n} F_i - \Phi(\mu, \sigma^2) \Big\|_K \leqslant \frac{C_1 \sum_{i=1}^{n} \beta_{3i}}{\left(\sum_{i=1}^{n} \sigma_i^2\right)^{3/2}}. \tag{2}$$

Here β_{3i} and σ_i^2 are the third absolute moment and variance of F_i, respectively. In many cases, the accuracy in (2) is of the order $O(n^{-1/2})$. However, this is not the case when random variables form triangular array and are close to zero.

Hipp [10] considered smoothing effect in general case of nonnegative random variables with some probability mass at zero. Here we present one improvement of Hipp's result by Roos, which follows from the more general proposition in [16]. Let all B_i be concentrated on $(0, \infty)$ and all $p_i < 1$, then

$$\Big\| \prod_{i=1}^{n} \big((1 - p_i)I + p_i B_i\big) - \exp\Big\{\sum_{i=1}^{n} p_i(B_i - I)\Big\} \Big\|_K \leqslant \frac{\pi^2}{4} \sum_{i=1}^{n} \frac{p_i^2}{1 - p_i} Q(\widetilde{H}, \mu_i). \tag{3}$$

Here $\mu_i = \int x \mathrm{d}B_i(x)$ and $\widetilde{H} = \exp\Big\{\sum_{i=1}^{n} p_i(1 - p_i)(B_i - I)/2\Big\}$. The smoothing effect is estimated by $Q(\widetilde{H}, \mu_i)$. Note that estimate without smoothing effect is equal to $C \min(\sum_1^n p_i^2, \max_i p_i)$, see Theorem 2.1, p.97 in [1].

Apart from the accompanying compound Poisson distribution as in (3) we consider the second order (signed) compound Poisson approximations, such as

$$\exp\Big\{\sum_{i=1}^{n}\Big(p_i(B_i-I)-p_i^2(B_i-I)^2/2\Big)\Big\}.$$

Analogues of (3) have been obtained for this approximation in [16]. For similar approximations see [2], [3], [15], and the references therein.

Note that lower bound estimates of compound Poisson approximation to weighted sums have been investigated in [5].

3 Results

1. Sums of 1-dependent random variables. First we consider the case, when random variables are non-identically distributed, that is, $S=w_1S_1+w_2S_2+\cdots+w_NS_N$ and

$$S_m=\sum_{i=1}^{n_m}X_{mi},\quad m=1,\dots,N.$$

We assume that S_m and S_j are independent when $m\neq j$. On the other hand, we allow weak dependence of variables in each sum. Let $X_{m1},X_{m2},\dots,X_{mn_m}$ be 1-dependent. We recall that the sequence of random variables $\{X_j\}_{j\geq 1}$ is called k-dependent if, for $1<s<t<\infty$, $t-s>m$, the sigma algebras generated by $X_1,\dots,X_s$ and $X_t,X_{t+1}\dots$ are independent. Though further on we consider 1-dependent variables, it is clear that, by grouping consecutive summands, the sum of k-dependent variables can be reduced to the sum of 1-dependent ones.

We consider the case when all X_{mk} are concentrated at $0,1,2\dots$. Factorial moments of X_{mk} are defined by

$$\nu_j^{(m)}(k)=\mathrm{E}X_{mk}(X_{mk}-1)\cdots(X_{mk}-j+1),\quad j=1,2,\dots,\quad m=1,\dots,N,\quad k=1,\dots,n_m.$$

Let

$$\Gamma_{m1}=\sum_{k=1}^{n_m}\nu_1^{(m)}(k),\quad \Gamma_{m2}=\frac{1}{2}\sum_{k=1}^{n_m}[\nu_2^{(m)}(k)-(\nu_1^{(m)}(k))^2]+\sum_{k=2}^{n_m}\mathrm{Cov}(X_{m,k-1},X_{mk}).$$

The distribution of w_mS_m we denote by F_m. Next we define approximating measures:

$$
\begin{aligned}
\Pi_m &= \exp\{\Gamma_{m1}(I_{w_m}-I)\}, \quad \Pi=\prod_{m=1}^N \Pi_m=\exp\Big\{\sum_{m=1}^N \Gamma_{m1}(I_{w_m}-I)\Big\},\\
G_m &= \exp\{\Gamma_{m1}(I_{w_m}-I)+\Gamma_{m2}(I_{w_m}-I)^2\}, \quad G=\prod_{m=1}^N G_m,\\
M_1 &= \exp\Big\{0.025\sum_{m=1}^N \Gamma_{m1}(I_{w_m}+I_{-w_m}-2I)\Big\}.
\end{aligned}
$$

Finally, we define remainder terms. Let $\widehat{\mathrm{E}}^+(Y_1,Y_2)=\mathrm{E}Y_1Y_2+\mathrm{E}Y_1\mathrm{E}Y_2$ and

$$
\begin{aligned}
R_{m0} &= \sum_{k=1}^n \Big\{\nu_2^{(m)}(k)+(\nu_1^{(m)}(k))^2+\mathrm{E}X_{m,k-1}X_{mk}\Big\},\\
R_{m1} &= \sum_{k=1}^n \Big\{(\nu_1^{(m)}(k))^3+\nu_1^{(m)}(k)\nu_2^{(m)}(k)+\nu_3^{(m)}(k)\\
&\quad +[\nu_1^{(m)}(k-2)+\nu_1^{(m)}(k-1)+\nu_1^{(m)}(k)]\mathrm{E}X_{m,k-1}X_{mk}\\
&\quad +\widehat{\mathrm{E}}^+(X_{m,k-1}(X_{m,k-1}-1),X_{mk})+\widehat{\mathrm{E}}^+(X_{m,k-1},X_{mk}(X_{mk}-1))\\
&\quad +\mathrm{E}X_{m,k-2}X_{m,k-1}X_{mk}+\mathrm{E}X_{m,k-2}\mathrm{E}X_{m,k-1}X_{m,k}\\
&\quad +\widehat{\mathrm{E}}^+(X_{m,k-2},X_{m,k-1})\mathrm{E}X_{mk}\Big\}.
\end{aligned}
$$

Theorem 3.1 *Let, for $m=1,2,\dots,N$; $k=1,2,\dots,n_m$, $\nu_1^{(m)}(k)\leqslant 1/100$, $\nu_2^{(m)}(k)\leqslant \nu_1^{(m)}(k)$, $\nu_3^{(m)}(k)<\infty$ and*

$$
\sum_{k=1}^{n_m}\nu_2^{(m)}(k)\leqslant\frac{\Gamma_{m1}}{20},\qquad \sum_{k=2}^{n_m}|\mathrm{Cov}(X_{m,k-1},X_{mk})|\leqslant\frac{\Gamma_{m1}}{20}. \tag{4}
$$

Then, for any $h>0$,

$$
\|F-\Pi\|_K \leqslant C_2Q(M_1,h)\sum_{m=1}^N R_{m0}\Big\{\frac{w_m}{h}\min\big(1,\Gamma_{m1}^{-1/2}\big)+\min\big(1,\Gamma_1^{-1}\big)\Big\}, \tag{5}
$$

$$
\|F-G\|_K \leqslant C_3Q(M_1,h)\sum_{m=1}^N R_{m1}\Big\{\frac{w_m}{h}\min\big(1,\Gamma_{m1}^{-1}\big)+\min\big(1,\Gamma_1^{-3/2}\big)\Big\}. \tag{6}
$$

Remark 3.1 *The choice of approximation in (6) is by no means restricted to G. For example, let $\Gamma_{m2}>0$. Then, taking into account Theorem 3.5 and corresponding Lemmas from [7] , it is* *possible reformulate (6) for the negative binomial approximation.*

As an application of Theorem 3.1 let us consider weighted sums of 2- runs. Two-runs statistic and its generalization k-runs statistic are one of the best investigated cases of sums of weakly dependent discrete random variables, see [3], [4], [8], [13], [19] and the references therein. Let $X_{mi}=\eta_{mi}\eta_{m,i+1}$, where $\eta_{mi}\sim Be(p_m)$, $(i=1,2,\dots,n_m+1)$ are independent Bernoulli variables. Then S_m is the sum of 1-dependent Bernoulli random variables. It is known that, if $n_m\geqslant 3$, $p_m\leqslant 1/5$, then

$$\begin{aligned}
\Gamma_{m1} &= np_m^2, \quad \Gamma_{m2}=\frac{n_mp_m^3(2-3p_m)-2p_m^3(1-p_m)}{2}, \quad R_{m1}\leqslant Cn_mp_m^4,\\
&\| F_m-G_m\|_K\leqslant \| F_m-G_m\|\leqslant C\frac{p_m}{\sqrt{n_m}}, \qquad (7)
\end{aligned}$$

see [13]. Therefore, the standard application of the triangle inequality as in (1) leads to estimate

$$\| F-G\|_K\leqslant C\sum_{m=1}^N\frac{p_m}{\sqrt{n_m}}. \qquad (8)$$

Let us assume that $w_m\asymp C$. If all p_m are sufficiently small, then conditions of Theorem 3.1 are satisfied. Therefore, taking $h=\min w_m/2$ in (18), we obtain

$$\| F-G\|_K\leqslant CQ(M_1,h)\sum_{m=1}^N p_m^2\leqslant\frac{C\sum_{m=1}^N p_m^2}{\sqrt{\sum_{m=1}^N n_mp_m^2}}. \qquad (9)$$

Estimate (9) can be much smaller than (8). If $p_i=p$, then the smoothing effect is very obvious:

$$\| F-G\|_K\leqslant\frac{C(N)p}{\sqrt{n_1+n_2+\dots+n_N}}\quad\text{vs}\quad \| F-G\|_K\leqslant C(N)p\bigg(\frac{1}{\sqrt{n_1}}+\dots+\frac{1}{\sqrt{n_N}}\bigg).$$

Note that due to 1-dependence we can not apply (3).

2. Sums of independent random variables satisfying Franken's condition. Theorem's 3.1 conditions can be relaxed if all random variables are independent. Let us consider typical case of clustered sample assuming that, in each sum, all random variables are independent and identically distributed. More precisely, let, for $m=1,2,\dots,N$, H_m be concentrated on lattice $0,w_m,2w_m,\dots,$

that is, $H_m = p_{m0}I + p_{m1}I_{w_m} + p_{m2}I_{2w_j} + \dots$. We denote jth factorial moment of H_m by

$$\nu_j(m) = \sum_{k=0}^{\infty} k(k-1)\cdots(k-j+1)p_{mk}$$

and assume Franken's condition

$$\lambda_m := \nu_1(m) - \nu_1^2(m) - \nu_2(m) > 0. \tag{10}$$

Franken [9] proved that, if the main probabilistic mass of nonnegative integer-valued random variable is concentrated at zero and unity, then the distribution of sum of such variables can be approximated by Poisson distribution quite accurately (see also [11]). Franken's condition means that $\nu_1(m) \leqslant 1$ and $\nu_2(m) \leqslant \nu_1(m)$. It is much weaker than Theorem's 3.1 assumptions $\nu_1(m) \leqslant 1/100$, $\nu_2(m) \leqslant \nu_1(m)$ and (4).

Theorem 3.2 *Let $\nu_3(m) < \infty$, $n_m \in \mathbb{N}$ and let condition (10) be satisfied,* $(m = 1, 2, \dots, N)$. *Then, for all* $h > 0$,

$$\Big\| \prod_{m=1}^{N} H_m^{n_m} - \exp\Big\{ \sum_{m=1}^{N} n_m \nu_1(m)\big(I_{w_m} - I\big) \Big\} \Big\|_K \leqslant C_4 Q(M_2, h)$$
$$\times \sum_{m=1}^{N} n_m(\nu_2(m) + \nu_1^2(m)) \Big\{ \frac{w_m}{h} \min\Big(1, \frac{1}{\sqrt{n_m\lambda_m}}\Big) + \min\Big(1, \frac{1}{n_m\lambda_m}\Big)\Big(1 + \frac{\nu_1(m)}{\lambda_m}\Big) \Big\} \tag{11}$$

and

$$\Big\| \prod_{m=1}^{N} H_m^{n_m} - \exp\Big\{ \sum_{m=1}^{N} n_m \Big(\nu_1(m)\big(I_{w_m} - I\big) + \frac{\nu_2(m) - \nu_1^2(m)}{2}(I_{w_m} - I)^2 \Big) \Big\} \Big\|_K$$
$$\leqslant C_5 Q(M_2, h) \sum_{m=1}^{N} n_m[\nu_3(m) + \nu_1(m)\nu_2(m) + \nu_1^3(m)]$$
$$\times \Big\{ \frac{w_j}{h} \min\Big(1, \frac{1}{n_m\lambda_m}\Big) + \min\Big(1, \frac{1}{(n_m\lambda_m)^{3/2}}\Big)\Big(1 + \frac{\nu_1(m)}{\lambda_m}\Big) \Big\}. \tag{12}$$

Here M_2 is symmetric distribution with $\widehat{M}_2(t) = \exp\Big\{ -\sum_{l=1}^{N} n_l\lambda_l \sin^2(tw_l/2) \Big\}$.

For any Bernoulli variable Franken's condition is satisfied. Therefore, assuming $h = \min_m w_m/2$ and applying (18), we obtain the following corollary.

Corollary 3.1 *Let $H_m = (1-p_m)I + p_m I_{w_m}$, $w_m \asymp C$, $p_m \leqslant C_6 < 1$, $m = 1, \dots, N$. Then*

$$\Big\| \prod_{m=1}^{N} H_m^{n_m} - \exp\Big\{ \sum_{m=1}^{N} n_m p_m \big(I_{w_j} - I\big) \Big\} \Big\|_K \leqslant C \Big(\sum_{i=1}^{N} n_i p_i \Big)^{-1/2} \sum_{m=1}^{N} \min\{n_m p_m^2, \sqrt{n_m} p_m^{3/2}\}. \quad (13)$$

It is easy to check, that if $N = n$, $n_j = 1$, then up to constant we get the classical estimate of Poisson approximation to the Poisson-binomial distribution with "magic factor" : $C \sum_1^n p_m^2 \Big(\sum_1^n p_j \Big)^{-1/2}$.

We also can use (13) for comparison to various known estimates. Let, in Corollary 3.1, $N = 2$ and $n_1 p_1 \geqslant 1$, $n_2 p_2 \geqslant 1$. Then the estimates in (2), (3) and (13) are of the order

$$\frac{1}{\sqrt{n_1 p_1 + n_2 p_2}}, \quad \frac{n_1 p_1^2 + n_2 p_2^2}{\sqrt{n_1 p_1 + n_2 p_2}}, \quad \frac{p_1 \sqrt{n_1 p_1} + p_2 \sqrt{n_2 p_2}}{\sqrt{n_1 p_1 + n_2 p_2}},$$

respectively. Here we used (18) for upper bound estimate in (3). It is easy to check, that the last estimate always has better order than the second one. Moreover, if p_1 and p_2 tend to zero sufficiently fast, the last estimate is sharper than the Berry-Esseen estimate.

3. Generalized Poisson-binomial distribution. We further relax assumptions on the structure of random variables and consider the case when all random variables are independent and have some probability mass at zero. The supports of random variables are unnecessary discrete and they might not have any finite absolute moment apart from the first one. We assume that random variables in each sum are identically distributed. In principle, we consider the case similar to the one considered in (3). However, we take an advantage of the fact that not all distributions are different. Let $\mu_{m1} = \int_{\mathbb{R}} |x| B_m\{dx\}$ and let $Re\widehat{B}_m(t)$ denote the real part of $\widehat{B}_m(t)$.

Theorem 3.3 *Let $B_j \in \mathcal{F}$, $0 \leqslant p_j \leqslant \tilde{C}_7 < 1$, $\mu_{m1} < \infty$ $(j = 1, \dots, N)$. Then, for any $h > 0$,*

$$\begin{aligned} &\Big\| \prod_{m=1}^{N} ((1-p_m)I + p_m B_m)^{n_m} - \exp\Big\{ \sum_{m=1}^{N} n_m p_m (B_m - I) \Big\} \Big\|_K \\ &\qquad \leqslant \quad C_8 Q(M_3, h) \sum_{m=1}^{N} n_m p_m^2 \Big\{ \frac{\mu_{m1}}{h} \min\left(1, \frac{1}{\sqrt{n_m p_m}}\right) + \min\left(1, \frac{1}{n_m p_m}\right) \Big\} \qquad (14) \end{aligned}$$

and

$$\Big\| \prod_{m=1}^{N}((1-p_m)I+p_mB_m)^{n_m} - \exp\Big\{\sum_{m=1}^{N}\big(n_mp_m(B_m-I)-\frac{n_m}{2}p_m^2(B_m-I)^2\big)\Big\}\Big\|_K$$
$$\leqslant\ C_9Q(M_3,h)\sum_{m=1}^{N}n_mp_m^3\Big\{\frac{\mu_{m1}}{h}\min\Big(1,\frac{1}{n_mp_m}\Big)+\min\Big(1,\frac{1}{(n_mp_m)^{3/2}}\Big)\Big\}. \qquad (15)$$

Here M_3 is symmetric distribution with $\widehat{M}_3(t)=\exp\Big\{\sum_{l=1}^{N}0.5n_lp_l(1-p_l)\big(Re\widehat{B}_l(t)-1\big)\Big\}$.

Remark 3.2 *(i) Though the accuracy of approximation is similar to that of previous Theorems, the structure of approximating Compound Poisson distribution is much more complicated.*

(ii) If, $B_m\{[0,\infty)\}=1$, $\mu_{m1}\asymp C$, $(m=1,2,\dots,N)$ then by (18) we can obtain estimate similar to (13). Therefore, it is not difficult to construct examples similar to the ones, considered for the previous theorem, and demonstrating the effect of smoothing.

(iii) If $n_j=1$, $N=n$, then (14) is a version of (3) for $B_m\{\mathbb{R}\}=1$. On the other hand, if all $B_m\{[0,\infty)\}=1$, then (3) is more accurate than (14).

4 Auxiliary results

Further we need the following lemmas.

Lemma 4.1 *Let $F,G\in\mathcal{F}$, $h>0$ and $a>0$. Then*

$$\begin{aligned} Q(F,h) &\leqslant \left(\frac{96}{95}\right)^2 h\int_{|t|\leqslant 1/h}\Big|\widehat{F}(t)\Big|\,\mathrm{d}t, && (16)\\ Q(F,h) &\leqslant \left(1+\left(\frac{h}{a}\right)\right)Q(F,a), && (17)\\ Q(\exp\{a(F-I)\},h) &\leqslant \frac{C}{\sqrt{aF\{|x|>h\}}}. && (18) \end{aligned}$$

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

$$h\int_{|t|\leqslant 1/h}|\widehat{F}(t)|\,\mathrm{d}t\leqslant CQ(F,h). \qquad (19)$$

Lemma 4.1 contains well-known properties of Levy's concentration function (see, for example, [1], Chapter 2).

For $h \in (0, \infty)$ and a finite nonnegative measure G on $\mathbb{R}$, set $|G|_{h-} = \sup_{x \in \mathbb{R}} G\{(x, x+h)\}$.

Lemma 4.2 *([6]) Let $W_1, W_2 \in \mathcal{M}$ with $W_1\{\mathbb{R}\} = 0$, and set $W = W_1 + W_2$. For $y \in [0, \infty)$, let*

$$\rho(y) = \min\big\{|W^+|_{y-},\, |W^-|_{y-}\big\}.$$

Then, for arbitrary $h \in (0, \infty)$ and $r \in (0, 1)$, we have

$$\| W \|_K \le \frac{1}{2r} \| W_1 \| + \frac{1}{2\pi r} \int_{|t|<1/h} \Big| \frac{\widehat{W}_2(t)}{t} \Big| \, \mathrm{d}t + \frac{1+r}{2r} \, \rho(4\, \eta(r) h),$$

where $\eta(r) \in (0, \infty)$ is defined by the equation

$$\frac{1+r}{2} = \frac{2}{\pi} \int_0^{\eta(r)} \frac{\sin^2(x)}{x^2} \, \mathrm{d}x.$$

Lemma 4.3 *([6]) For $F \in \mathcal{F}$, $W \in \mathcal{M}$ with $W\{\mathbb{R}\} = 0$, and $\vartheta \in (0, \infty)$, we have*

$$|(WF)^+|_{\vartheta-} \le \frac{1}{2} \| W \| \, |F|_{\vartheta-}. \tag{20}$$

From Lemmas 4.2 and 4.3 and (16) and (19) the following result follows

Lemma 4.4 *Let $h > 0$, $W \in \mathcal{M}$, $W\{\mathbb{R}\} = 0$, $P \in \mathcal{F}$, M be distribution with nonnegative characteristic function and $|\widehat{P}(t)| \leqslant C M(t)$, for $|t| \leqslant 1/h$. Then*

$$\begin{aligned} \| WP \|_K &\leqslant C \int_{|t| \leqslant 1/h} \Big| \frac{\widehat{W}(t) \widehat{P}(t)}{t} \Big| \, \mathrm{d}t + C \, \|W\| \, Q(P, h) \\ &\leqslant C \Big(\sup_{|t| \leqslant 1/h} \frac{|\widehat{W}(t)|}{|t|} \cdot \frac{1}{h} + \|W\| \Big) Q(M, h). \end{aligned}$$

Proof. We apply Lemma 4.2 with $W_1 = 0$, $W = W_2 = WP$, and $r = 0.5$. Then by (17) and (20) we have

$$\rho(4\, \eta(r) h) \leqslant C |(WP)^+|_{4\eta(r)h} \leqslant \| W \| Q(P, 4\eta(r)h) \leqslant C \| W \| Q(P, h).$$

Moreover, applying (16) and (19), we prove that

$$Q(P,h) \leqslant Ch\int_{|t|\leqslant 1/h} |\widehat{P}(t)|\,\mathrm{d}t \leqslant Ch\int_{|t|\leqslant 1/h} \widehat{M}(t)\,\mathrm{d}t \leqslant CQ(M,h)$$

and

$$\int_{|t|\leqslant 1/h} |\widehat{P}(t)|\,\mathrm{d}t \leqslant C\frac{1}{h}h\int_{|t|\leqslant 1/h} \widehat{M}(t)\,\mathrm{d}t \leqslant \frac{1}{h}CQ(M,h).$$

This, obviously, completes the proof of Lemma. □

Lemma 4.5 *Let $M \in \mathcal{F}$ be concentrated on integers, $\sum_{k=-\infty}^{\infty} |kM\{k\}| < \infty$. Then, for all $\gamma > 0$ and $v \in R$,*

$$\| M \|^2 \leqslant \Big(\frac{1}{2} + \frac{1}{2\pi\gamma}\Big)\int_{-\pi}^{\pi}\Big(\gamma|\widehat{M}(t)|^2 + \frac{1}{\gamma}\Big|\big(\widehat{M}(t)e^{-itv}\big)'\Big|^2\Big)\,\mathrm{d}t.$$

Lemma 4.5 has been proved in [14].

Lemma 4.6 *Let conditions of Theorem 3.1 be satisfied. Then, for all $t \in \mathbb{R}$, $m = 1,\dots,N$,*

$$
\begin{aligned}
|\widehat{F}_m(t)|, |\widehat{G}_m(t)|, |\widehat{\Pi}_m(t)| &\leqslant \exp\{-0.26\Gamma_{m1}\sin^2(t_m/2)\}, &(21)\\
|\widehat{F}_m(t) - \widehat{G}_m(t)| &\leqslant CR_{m1}|z_m(t_m)|^3\psi_m^{2.6}, &(22)\\
\Big|\Big(\exp\{-\mathrm{i}t_m\Gamma_{m1}\}(\widehat{F}_m(t) - \widehat{G}_m(t))\Big)'_{t_m}\Big| &\leqslant CR_{m1}|z(t_m)|^2(1 + |z(t_m)|^2\Gamma_{m1})\psi_m^{2.6} \\
&\leqslant CR_{m1}|z(t_m)|^2\psi_m^2, &(23)\\
|\widehat{F}_m(t) - \widehat{\Pi}(t)| &\leqslant CR_{m0}|z(t_m)|^2\psi_m^{2.6}, &(24)\\
\Big|\Big(\exp\{-\mathrm{i}t_m\Gamma_{m1}\}(\widehat{F}_m(t) - \widehat{\Pi}_m(t))\Big)'_{t_m}\Big| &\leqslant CR_{m0}|z(t_m)|\psi_m^2. &(25)
\end{aligned}
$$

Here $t_m = tw_m$, $z(t_m) = \mathrm{e}^{\mathrm{i}t_m} - 1$, $\psi_m = \exp\{-0.1\Gamma_{m1}\sin^2(tw_m/2)\}$.

All estimates in Lemma 4.6 follow from Lemmas 7.4, 7.6, 7.7 and the proof of theorem 5.1 in [7].

5 Proofs

As in previous Section $z(t) = \mathrm{e}^{\mathrm{i}t} - 1$, $t_m = tw_m$, $\psi_m = \exp\{-0.1\Gamma_{m1}\sin^2(tw_m/2)\}$. We use the notation θ for all quantities satisfying $|\theta| \leqslant 1$.

Proof of Theorem 3.1. By properties of the total variation norm

$$\begin{aligned}
\| F-G \|_K &= \Big\| \prod_{m=1}^{N} F_m - \prod_{m=1}^{N} G_m \Big\|_K \leqslant \sum_{m=1}^{N} \Big\| (F_m-G_m) \prod_{l=1}^{m-1} F_l \prod_{l=m+1}^{N} G_l \Big\|_K \\
&= \Big\| (F_m-G_m)\exp\{-0.05\Gamma_{m1}(I_{w_m}-I)\} \Big\|_K \\
&\quad \times \Big\| \exp\{0.05\Gamma_{m1}(I_{w_m}-I)\} \prod_{l=1}^{m-1} F_l \prod_{l=m+1}^{N} G_l \Big\|_K \\
&=: \sum_{m=1}^{N} \| W_m P_m \|_K. \qquad (26)
\end{aligned}$$

Note that $\exp\{-0.05\Gamma_{m1}(I_{w_m}-I)\}$ is signed measure of finite variation.

Applying (21) we obtain

$$|\widehat{P}_m(t)| \leqslant C\psi_m \prod_{l\neq m}^{N} \psi_l^{2.6} \leqslant M_1(t).$$

Similarly, from (22) it follows that

$$|W(t)| \leqslant CR_{m1}|z(t_m)|^3\psi_m^{2.6}\psi_m^{-1} \leqslant CR_{m1}|z(t_m)|^2 w_m|t|\psi_m^{1.6} \leqslant CR_{m1}\min(1,\Gamma_{m1}^{-1})\psi_m^{0.5}w_m|t|.$$

Here $\psi_m = \exp\{-0.1\Gamma_{m1}\sin^2(tw_m/2)\}$. Applying Lemma 4.4 we obtain

$$\| W_m P_m \|_K \leqslant CQ(M_1,h)\Big\{ R_{m1}\min\big(1,\Gamma_{m1}^{-1}\big)\frac{w_m}{h} + \| W_m \| \Big\}. \qquad (27)$$

It remains to estimate $\| W_m \|$. Since, total variation norm is invariant to scale change, further we assume $w_m = 1$, $t_m = t$. Then, applying Lemma (4.6), we obtain

$$\begin{aligned}
|\widehat{W}_m(t)| &\leqslant CR_{m1}\min(1.\Gamma_{m1}^{-3/2})\psi_m^2, \\
\Big|\Big(\exp\{-0.9\mathrm{i}t\Gamma_{m1}\}\widehat{W}_m(t)\Big)_t'\Big| &\leqslant \Big|\Big(\exp\{-\mathrm{i}t\Gamma_{m1}\}(\widehat{F}_m(t)-\widehat{G}_m(t))\Big)_t'\Big| \exp\{0.1\Gamma_{m1}\sin^2(t/2)\} \\
&\quad + C|\widehat{F}_m(t)-\widehat{G}_m(t)|\Gamma_{m1}|\sin(t/2)|\exp\{0.1\Gamma_{m1}\sin^2(t/2)\} \\
&\leqslant CR_{m1}\sin^2(t/2)\psi_m^{1.5}(1+\Gamma_{m1}\sin^2(t/2)) \leqslant CR_{m1}\min(1,\Gamma_{m1}^{-1})\psi_m.
\end{aligned}$$

Taking into account the last two estimates and, applying Lemma 4.5 with $\gamma = \max(1,\sqrt{\Gamma_{m1}})$,

$\upsilon = 0.9\Gamma_{m1}$, we get

$$\| W_m \| \leqslant C R_{m1} \min(1, \Gamma_{m1}^{-3/2}).$$

Substituting the last estimate estimate into (27) and (26) we complete the proof of (6). The proof of (5) is very similar and, therefore, omitted. $\square$

Proof of Theorem 3.2. The estimates are proved exactly by the same arguing as in the proof of Theorem 3.1. Let

$$\begin{aligned}
\widehat{D}_m(t) &= \exp\Big\{\nu_1(m)z(t_m) + (\nu_2(m) - \nu_1^2(m))\frac{z^2(t_m)}{2}\Big\},\\
\widehat{V}_m(t) &= \big(\widehat{H}_m^{n_m}(t) - \widehat{D}_m^{n_m}(t)\big)\exp\{-0.5 n_m \lambda_m z(t_m)\},\\
L_m &= \exp\Big\{0.5\lambda_m(I_{w_m} - I)\Big\}\prod_{l=1}^{m-1} H_l^{n_l}\prod_{l=m+1}^{N} D_l^{n_l}.
\end{aligned}$$

Then

$$\Big\| \prod_{m=1}^{N} H_m^{n_m} - \prod_{m=1}^{N} D_m^{n_m} \Big\|_K \leqslant \sum_{m=1}^{N} \| V_m L_m \|_K. \tag{28}$$

Distribution H_m and approximation D_m, for the case $w_m = 1$, were investigated in numerous papers. Let $r_1(m) = \nu_1^3(m) + \nu_1(m)\nu_2(m) + \nu_3(m)$. Taking into account Lemmas 2 and 3 in [18] and proof of Theorem 3 in [11] we can write the following expressions

$$\begin{aligned}
\widehat{H}_m(t) &= 1 + \nu_1(m)z(t_m) + \frac{\nu_2(m)}{2}z^2(t_m) + \theta C\nu_3(m)|z(t_m)|^3,\\
\big(\widehat{H}_m(t)\big)'_{t_m} &= 1 + \nu_1(m)(z(t_m))' + \frac{\nu_2(m)}{2}(z^2(t_m))' + \theta C\nu_3(m)|z(t_m)|^2,
\end{aligned}$$

$$\begin{aligned}
|\widehat{H}_m(t)|,\ |\widehat{D}_m(t)|, |\exp\{\nu_1(m)z(t_m)\}| &\leqslant \exp\{-2\lambda_m\sin^2(t_m/2)\},\\
|\widehat{H}_m(t) - \exp\{\nu_1(m)z(t_m)\}| &\leqslant C(\nu_1^2(m) + \nu_2(m))|z(t_m)|^2,\\
|(\widehat{H}_m(t) - \exp\{\nu_1(m)z(t_m)\})'_{t_m}| &\leqslant C(\nu_1^2(m) + \nu_2(m))|z(t_m)|,\\
|\widehat{H}_m(t) - \widehat{D}_m(t)| &\leqslant C r_1(m)|z(t_m)|^3,\\
|(\widehat{H}_m(t) - \widehat{D}_m(t))'_{t_m}| &\leqslant C r_1(m)|z(t_m)|^2.
\end{aligned}$$

Therefore, $|L_m(t)| \leqslant C\widehat{M_2}(t)$, and

$$\begin{aligned}
|\widehat{V}_m(t)| &\leqslant Cn_m|\widehat{H}_m(t) - \widehat{D}_m(t)| \exp\{n_m\lambda_m \sin^2(t_m/2) - 2(n_m-1)\lambda_m \sin^2(t_m/2)\} \\
&\leqslant C\exp\{-n_m\lambda_m \sin^2(t_m/2)\} r_1(m)|z(t_m)|^3 \\
&\leqslant C\exp\{-0.5n_m\lambda_m \sin^2(t_m/2)\} r_1(m) \min(1, (n_m\lambda_m)^{-1}) w_m|t|.
\end{aligned}$$

Applying Lemma 4.4 we obtain

$$\| V_mL_m \|_K \leqslant CQ(M_2, h)\Big\{ n_m r_1(m) \min\big(1, (n_m\lambda_m)^{-1}\big) \frac{w_m}{h} + \| V_m \| \Big\}. \tag{29}$$

It remains to estimate $\| V_m \|$. Total variation norm is invariant to scale change. Therefore, we can assume $w_m = 1$, $t_m = t$. For the sake of brevity we use the following notation omitting the dependence on t and m:

$$\omega = \exp\{-0.5n_m\lambda_m \sin^2(t/2)\}, \quad u_1 = \widehat{H}(t)\exp\{-\mathrm{i}t\nu_1(m)\}, \quad u_2 = \widehat{D}(t)\exp\{-\mathrm{i}t\nu_1(m)\}.$$

Taking into account relations from above, we can write $|\widehat{V}_m(t)| \leqslant \omega^2 r_1(m)\min(1, (n_m\lambda_m)^{-3/2})$ and

$$\begin{aligned}
|(\widehat{V}_m(t)\exp\{-n_m\nu_1(m)\mathrm{i}t + 0.5n_m\lambda_m\mathrm{i}t\})'| &= |\big((u_1^{n_m} - u_2^{n_m})\exp\{0.5n_m\lambda_m(1 + \mathrm{i}t - \mathrm{e}^{\mathrm{i}t})\}\big)'| \\
&\leqslant n_m|u_1^{n_m-1}u_1' - u_2^{n_m-1}u_2'|\omega^{-2} + |u_1^{n_m} - u_2^{n_m}|0.5n_m\lambda_m|z(t)|\omega^{-2} \\
&\leqslant n_m|u_1|^{n_m-1}|u_1' - u_2'|\omega^{-2} + n_m|u_2'||u_1^{n_m-1} - u_2^{n_m}|\omega^{-2} \\
&\quad + |u_1^{n_m} - u_2^{n_m}|0.5n_m\lambda_m|z(t)|\omega^{-2} \\
&\leqslant n_m|u_1' - u_2'|\omega^2 + Cn_m^2(\nu_2 + \nu_1)r_1(m)|z(t)|^4\omega^2 + Cn_m^2\lambda_m r_1(m)|z(t)|^4\omega^2 \\
&\leqslant Cn_m r_1(m)|z(t)|^2\omega^{1.5}\big(1 + n_m[\nu_1(m) + \lambda_m]|z(t)|^2\omega^{0.5}\big) \\
&\leqslant Cn_m r_1(m)\omega\left(1 + \frac{\nu_1(m)}{\lambda_m}\right)\min\left(1, \frac{1}{n_m\lambda_m}\right).
\end{aligned}$$

Applying Lemma 4.5 with $\gamma = \max(1, \sqrt{n_m\lambda_m})$, $\upsilon = n_m\nu_1(m) - 0.5n_m\lambda_m$, we get

$$\| V_m \| \leqslant Cn_m r_1(m)\left(1 + \frac{\nu_1(m)}{\lambda_m}\right)\min\left(1, \frac{1}{(n_m\lambda_m)^{3/2}}\right). \tag{30}$$

Combining the last estimate with (29) and (28) we complete the proof of (12). The proof of (11) is very similar and , therefore, omitted. □

Proof of Theorem 3.3. Similarly to the proof of previous Theorem we prove that

$$\Big\| \prod_{m=1}^{N} ((1-p_m)I + p_m B_m)^{n_m} - \exp\Big\{\sum_{m=1}^{N} p_m(B_m - I)\Big\}\Big\|_K \leqslant \sum_{m=1}^{N} \| U_m T_m \|_K.$$

Here

$$\begin{aligned}
\widehat{U}_m(t) &= \big[((1-p_m) + p_m\widehat{B}_m(t))^{n_m} - \exp\{n_m p_m(\widehat{B}(t)-1)\}\big] \\
&\quad \times \exp\{0.5 n_m p_m(1-p_m)(1-\widehat{B}_m(t))\},\\
\widehat{T}_m(t) &= \exp\{0.5 n_m p_m(1-p_m)(\widehat{B}_m(t)-1)\}\\
&\quad \times \prod_{j=1}^{m-1}((1-p_j)+p_j\widehat{B}_j(t))^{n_j} \prod_{j=m+1}^{N} \exp\{n_j p_j(\widehat{B}_j(t)-1)\}.
\end{aligned}$$

Taking into account general estimate, $|\widehat{B}_m(t)-1|^2 \leqslant 2|Re\widehat{B}_m(t)-1|$ we obtain:

$$\begin{aligned}
|\exp\{p_m(\widehat{B}_m(t)-1)\}| &= \exp\{p_m(Re\widehat{B}_m(t)-1)\},\\
|\exp\{p_m(\widehat{B}_m(t)-1) - 0.5p_m^2(\widehat{B}_m(t)-1)^2\}| &\leqslant \exp\{p_m(1-p_m)(Re\widehat{B}_m(t)-1)\},
\end{aligned}$$

$$|1+p_m(\widehat{B}_m(t)-1) - \exp\{p_m(\widehat{B}_m(t)-1)\}| \leqslant Cp_m^2|\widehat{B}_m(t)-1|^2 \leqslant Cp_m^2|Re\widehat{B}_m(t)-1|^{1/2}\mu_{m1}|t|,$$

$$|1+p_m(\widehat{B}_m(t)-1) - \exp\{p_m(\widehat{B}_m(t)-1) - 0.5p_m^2(\widehat{B}_m(t)-1)^2\}| \leqslant Cp_m^3|Re\widehat{B}_m(t)-1|\mu_{m1}|t|,$$

Consequently,

$$\begin{aligned}
|\widehat{T}_m(t)| &\leqslant \exp\Big\{0.5\sum_{m=1}^{N} n_m p_m(1-p_m)(Re\widehat{B}_m(t)-1)\Big\} = \widehat{M}_3(t), && (31)\\
|\widehat{U}_m(t)| &\leqslant Cn_m|1+p_m(\widehat{B}_m(t)-1) - \exp\{p_m(\widehat{B}_m(t)-1)\}|\exp\{0.5p_m(1-p_m)(Re\widehat{B}_m(t)-1)\}\\
&\leqslant Cn_m p_m^2\mu_{m1}|t|\min\Big(1, \frac{1}{\sqrt{n_m p_m}}\Big). && (32)
\end{aligned}$$

Moreover, due to the properties of total variation norm,

$$\| U_m \| \leqslant \| \big((I+p_m(I_1-I))^{n_m} - \exp\{n_m p_m(I_1-I)\}\big)\exp\{0.5n_m p_m(1-p_m)(I-I_1)\} \|.$$

Arguing similarly as in the proof of (30) we prove that

$$\| U_m \| \leqslant C n_m p_m^2 \min\left(1, \frac{1}{n_m p_m}\right). \tag{33}$$

From Lemma 4.4 and (31)–(33) we obtain (14). The proof of estimate (15) is very similar and , therefore, omitted. □

References

[1] Arak, T.V. and Zaĭtsev, A. Yu.(1988) Uniform limit theorems for sums of independent random variables. *Proc. Steklov Instit. Math.*, **174**(1).

[2] Barbour, A.D. and Čekanavičius, V. (2002) Total variation asymptotics for sums of independent integer random variables. *Ann. Probab.*, **30**(3), 509–545.

[3] Barbour, A.D. and Xia, A. (1999) Poisson perturbations. *ESAIM: Probab. Statist.*, **3**, 131–150.

[4] Brown, T.C. and Xia, A. (2001) Stein's method and birth-death processes, *Ann. Probab.,* **29**, 1373–1403.

[5] Čekanavičius, V. and Elijio, A. (2005) Lower-bound estimates for Poisson type approximations. *Lith. Math. J.*, **45**, 405–423.

[6] Čekanavičius, V. and Roos, B. (2006) Compound Binomial Approximations. *Ann. Inst. Statist. Math.*, **58**(1), 187–210 (2006).

[7] Čekanavičius, V. and Vellaisamy, P. (2013) Discrete approximations for sums of m-dependent random variables, (submitted for publication, preprint version is at arXiv:1301.7196).

[8] Daly,F., Lefevre,C. and Utev, S. (2012) Stein's method and stochastic orderings. *Adv. Appl. Prob.*, **44**, 343–372.

[9] Franken, P. (1964) Approximation des Verteilungen von Summen unabhängiger nichtnegativer ganzzähliger Zufallsgrößen durch Poissonsche Verteilungen. *Math. Nachr.*, **27**, 303–340.

[10] Hipp, C. (1985) Approximation of aggregate claims distributions by compound Poisson distributions. *Insurance*: *Math. and Economics.*, **4**, 227–232. Correction note: 6, 165 (1987).

[11] Kruopis, J. (1986) Approximations for distributions of sums of lattice random variables I. *Lith. Math. J.*, **26**, 234-244.

[12] Liang, Han-Ying and Baek, Jong-Il (2006) Weighted sums of negatively associated random variables. *Aust. N.Z. J. Stat.*, **48**(1), 21–31.

[13] Petrauskienė, J. and Čekanavičius, V. (2010) Compound Poisson approximations for sums of 1-dependent random variables I. *Lith. Math. J.,* **50**, 323–336.

[14] Presman, E.L. (1986) Approximation in variation of the distribution of a sum of independent Bernoulli variables with a Poisson law. *Theory Probab. Appl.*, **30**(2), 417–422.

[15] Roos, B. (2003) Poisson approximation via the convolution with Kornya-Presman signed measures. *Theory Probab. Appl.,* **48**, 555–560.

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

[17] Rosalsky, A. and Sreehari, M., (1998). On the limiting behavior of randomly weighted partial sums. *Statist. Probab. Lett.*, **40**, 403–410.

[18] Šiaulys, J. and Čekanavičius, V. (1988) Approximation of distributions of integer-valued additive functions by discrete charges I. *Lith. Math. J.*, **28**, 392–401.

[19] Wang, X. and Xia, A. (2008) On negative binomial approximation to k-runs. *J. Appl. Probab.,* **45**, 456–471.

[20] Zhang, Li-Xin (1997) Strong approximation theorems for geometrically weighted random series and their applications. *Ann. Probab.*, **25**(4), 1621–1635.