Bernoulli **10**(4), 2004, 665–683

On local estimates and the Stein method

VYDAS ČEKANAVIČIUS

Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, Vilnius 2600, Lithuania. E-mail: vydas.cekanavicius@maf.vu.lt

We show that recursive application of the Stein equation reduces uniform and non-uniform local estimates to estimates of the difference operator in total variation. We illustrate our approach by an example of an sth-order signed compound Poisson approximation to the sum of integer-valued random variables.

Keywords: local estimates; Stein method; total variation

1. Introduction

The elegant adaptation of the Stein method by Chen (1975) to the case of integer-valued random variables is one of the most powerful techniques in the theory of approximations. The method employs properties of the solution of a special difference equation. Usually, the Stein–Chen method is used for estimates in total variation; see, for example, Barbour *et al.* (1992a; 1992b), Barbour and Chryssaphinou (2001), Brown and Xia (2001), Barbour and Čekanavičius (2002), and references therein. The slightly modified Stein–Chen method can be used for moderately large deviations; see Chen and Choi (1992). As shown by Barbour and Jensen (1989), the method is also suitable for uniform local estimates.

Only a few papers are devoted to discrete non-uniform estimates; see Barbour (1987), Barbour *et al.* (1995), Chen and Roos (1995) and Chen (2000). Direct application of the Stein–Chen method requires solving the Stein equation for unbounded functions, which is a serious technical problem. So far, the most accurate estimates strongly depend on the properties of the approximating Poisson distribution and on the independence of random variables. Note that, for estimates in total variation, the Stein method is extended to cases more general than that of the Poisson approximation.

In this paper, we show how recursive application of the Stein equation can reduce non-uniform local estimates to estimates of a difference operator in total variation. Our approach has the following advantages: there is no need to solve the Stein equation for unbounded functions; for the numerous cases with already established estimates in total variation, local non-uniform estimates can be obtained without much additional effort; and, in principle, the proposed approach can be extended to the sums of dependent random variables. On the other hand, the non-uniform functions in this paper are growing no faster than polynomials. For a Poisson approximation, a similar recursive approach was used by Barbour (1987). Note that Barbour (1987) and Barbour *et al.* (1995) used more general unbounded functions.

The method of our paper is best suited to approximations depending on more than one parameter. That is, it is suited to compound Poisson approximations or asymptotic expansions, rather than to the standard Poisson approximation. We consider cases where a *perturbation* argument can be applied. Introduced by Barbour and Xia (1999), the perturbation argument simplifies the matter of solving the Stein equation for compound Poisson approximations. We illustrate our approach by constructing an s-parametric signed compound Poisson (SCP) approximation. SCP measures can be viewed as a special kind of compounding asymptotic expansion. Their main advantages are their infinite divisibility and compound Poisson structure. In its simplest form, an SCP measure can be expressed as the convolution of Poisson-like measures with possibly negative parameters. An SCP approximation is not a distribution but rather a signed measure. In this sense, SCP measures do not differ from other asymptotic expansions. On SCP approximations, see Kruopis (1986), Barbour and Xia (1999), Barbour and Čekanavičius (2002), and Roos (2002).

We need the following notation. Let E_a denote the distribution concentrated at a point a, $E \equiv E_0$. For a (signed) measure G, its total variation norm is denoted by $\|G\|$. If G is concentrated on $\mathbb{Z}$, then we write $\|G\| = \sum_j \|G(j)\|$. The convolution of measures G and F is denoted by $G * F$. If F and G are concentrated on the integers, then

$$F * G\{m\} = \sum_{k=-\infty}^{\infty} F\{m-k\}G\{k\}.$$

Throughout the paper, f denotes a function $f\colon \mathbb{Z}_+ \to \mathbb{R}$. We denote the supremum norm and ℓ_1 norm by $\|f\| = \sup_{j\geqslant 0}|f(j)|$ and $\|f\|_1 = \sum_0^\infty |f(j)|$, respectively. The difference operator Δ is defined by $\Delta f(j) = f(j+1) - f(j)$, $\Delta^k f(j) = \Delta^{k-1} f(j+1) - \Delta^{k-1} f(j)$, and $\Delta^0 f(j) = f(j)$.

Let Z_i, $i = 1, 2, \ldots, n$, be independent non-negative integer-valued random variables, and let

$$q_{ij} = P(Z_i = j), \qquad \mu_i = \mathrm{E}Z_i = \sum_{j=1}^{\infty} j q_{ij}, \qquad \sigma_i^2 = \mathrm{var}\, Z_i,$$

$$W = \sum_{i=1}^{n} Z_i, \qquad W^{(i)} = W - Z_i, \qquad W^{(ij)} = W - Z_i - Z_j,$$

$$\mu = \mathrm{E}W = \sum_{i=1}^{n} \mu_i, \qquad \sigma^2 = \mathrm{var}\, W = \sum_{i=1}^{n} \sigma_i^2,$$

$$\mathrm{E}(Z_i)_{(k)} = \mathrm{E}Z_i(Z_i - 1)\cdots(Z_i - k + 1), \qquad d_k = \|\mathcal{L}(W) * (E_1 - E)^{*k}\|,$$

$$d'_k = \sup_i \|\mathcal{L}(W^{(i)}) * (E_1 - E)^{*k}\|, \qquad d''_k = \sup_{i,j} \|\mathcal{L}(W^{(ij)}) * (E_1 - E)^{*k}\|.$$

Note that $\mathrm{E}(Z_i)_{(k)}$ is the kth factorial moment of Z_i. The quantities d_k, d'_k, and d''_k are used for approximating $\mathcal{L}(W)$ in total variation; see Barbour and Xia (1999) and Barbour and Čekanavičius (2002). Moreover, as proved by Barbour and Xia (1999),

$$d_1 = \|\mathcal{L}(W) * (E_1 - E)\| = \|\mathcal{L}(W+1) - \mathcal{L}(W)\| \leqslant 2V^{-1/2}, \tag{1.1}$$

where

$$V = \sum_{i=1}^{n} v_i, \qquad v_i = \min\{1/2;\ 1 - \|\mathcal{L}(Z_i + 1) - \mathcal{L}(Z_i)\|/2\}.$$

The properties of total variation and (1.1) allow one to obtain similar estimates for other d_ks. For example, let us assume that $V > 4$. Then we can write

$$W^{(i)} = W_1^{(i)} + W_2^{(i)} + W_3^{(i)},$$

where

$$W_j^{(i)} = \sum_{i \in S_j} Z_i.$$

Moreover, the sets of indices S_j can be chosen so that

$$\sum_{i \in S_j} v_i \geqslant (V - v^*)/3 - v^* = (V - 4v^*)/3, \qquad v^* = \max v_i.$$

Taking into account (1.1) and the properties of the total variation norm, we obtain

$$d_3' = \sup_i \|\mathcal{L}(W_1^{(i)}) \, * \, \mathcal{L}(W_2^{(i)}) \, * \, \mathcal{L}(W_3^{(i)}) \, * \, (E_1 - E)^{*3}\|$$

$$\leqslant \sup_i \prod_{j=1}^{3} \|\mathcal{L}(W_j^{(i)}) \, * \, (E_1 - E)\| \leqslant 24\sqrt{3}(V - 4v^*)^{-3/2}. \tag{1.2}$$

Similarly, $d_2' \leqslant 8(V - v^*)^{-1}$; see Barbour and Čekanavičius (2002). In general, d_k, d_k', and d_k'' are of order $O(V^{-k/2})$ as $V \to \infty$.

The structure of the paper is as follows. In Section 2, we establish properties of the solution of the Stein equation for SCP approximation and outline the basic idea of the recursive algorithm. Section 3 is devoted to auxiliary estimates of the difference operator. In Section 4, we formulate the main result. In Section 5, we discuss a possible extension to the case of dependent random indicators.

2. The Stein equation for SCP approximations

Let π denote the (possibly signed) measure with generating function

$$\hat{\pi}(z) = \sum_{j=0}^{\infty} \pi(j) z^j = \exp\left\{\sum_{l=1}^{\infty} \lambda_l (z^l - 1)\right\}, \qquad \lambda_l \in \mathbb{R}. \tag{2.1}$$

If all $\lambda_l \geqslant 0$, then π is a compound Poisson distribution, otherwise π is an SCP measure. Let us assume that

$$\lambda = \sum_{l=1}^{\infty} l\lambda_l, \qquad (Ug)(j) := \sum_{l\geqslant 2} l\lambda_l \sum_{k=1}^{l-1} \Delta g(j+k), \tag{2.2}$$

$$(\mathcal{A}g)(j) := \sum_{l\geqslant 1} l\lambda_l g(j+l) - jg(j) = \lambda g(j+1) - jg(j) + (Ug)(j). \tag{2.3}$$

$\mathcal{A}$ is called the Stein operator for π. It is defined on $g : \mathbb{Z}_+ \to \mathbb{R}$ and, for all bounded g, satisfies the equation

$$\pi\{\mathcal{A}g\} := \sum_{j=0}^{\infty} (\mathcal{A}g)(j)\pi(j) = 0. \tag{2.4}$$

If

$$\lambda > 0, \qquad \theta = \lambda^{-1} \sum_{l\geqslant 2} l(l-1)|\lambda_l| < 1/2, \tag{2.5}$$

then, for any bounded f, the solution of the Stein equation

$$(\mathcal{A}g)(j) = f(j) - \pi\{f\} \tag{2.6}$$

satisfies the inequalities

$$\|g\| \leqslant \frac{2\|f\|}{1-2\theta}(1 \wedge \lambda^{-1/2}), \qquad \|\Delta g\| \leqslant \frac{2\|f\|}{1-2\theta}(1 \wedge \lambda^{-1}); \tag{2.7}$$

see Barbour and Xia (1999).

We first note that, for $\|f\|_1 < \infty$, the solution of (2.6) has even better properties than those given by (2.7).

Lemma 2.1. *Let $\|f\|_1 < \infty$. If* (2.5) *is satisfied, then the solution of* (2.6) *has the following properties:*

$$g(i) = 0, \qquad i \leqslant 0;$$
$$\|\Delta g\|_1 \leqslant 2\|f\|_1 (1-2\theta)^{-1}\lambda^{-1};$$
$$\|g\| \leqslant 2\|f\|_1 (1-2\theta)^{-1}\lambda^{-1};$$
$$\|Ug\|_1 \leqslant 2\theta\|f\|_1 (1-2\theta)^{-1}\lambda^{-1}.$$

Expression (2.1) is related to expansions of generating functions in cumulants. An alternative approach involves using factorial cumulants. It is applied to the sum of random variables W defined in Section 1.

Let π_s be the SCP measure with the moment generating function

$$\hat{\pi}_s(z) = \exp\left\{\sum_{j=1}^{s} \beta_j (z-1)^j\right\}, \qquad \beta_j \in \mathbb{R},\ s \geqslant 1. \tag{2.8}$$

In (2.8), we assume that $\beta_j = \sum_{i=1}^n \beta_j^{(i)}$. Note that the jth factorial cumulant of Z_i equals $\beta_j^{(i)} j!$. We choose β_j for $\mathcal{L}(W)$ and π_s, ensuring the matching of s moments. Moreover,

$$\beta_1 = \mu = \sum_{i=1}^n \mu_i.$$

Note that π_1 corresponds to the standard Poisson approximation. Obviously, $\hat{\pi}_s(z)$ is the partial case of (2.1) where $\lambda = \beta_1 = \mu$,

$$\lambda_j = \sum_{k=j}^s \binom{k}{j} (-1)^{k-j} \beta_k, \qquad j = 1, \ldots, s; \tag{2.9}$$

and $\lambda_j = 0$ for $j > s$.

The Stein operator corresponding to π_s can be written in the form

$$(\mathcal{A}_s g)(j) := \beta_1 g(j+1) - jg(j) + \sum_{k=2}^s k\beta_k \Delta^{k-1} g(j+1). \tag{2.10}$$

Indeed, by (2.9) we have

$$\begin{aligned}\lambda_1 g(j+1) + \sum_{l=2}^\infty l\lambda_l g(j+l) &= \lambda_1 g(j+1) + \sum_{l=2}^s \sum_{k=l}^s l \binom{k}{l} (-1)^{k-l} g(j+l)\beta_k \\ &= \beta_1 g(j+1) + \sum_{k=2}^s k\beta_k \sum_{l=0}^{k-1} \binom{k-1}{l} (-1)^{k-l-1} g(j+1+l) \\ &= \sum_{k=1}^s k\beta_k \Delta^{k-1} g(j+1).\end{aligned}$$

Moreover,

$$(Ug)(j) = \sum_{k=2}^s k\beta_k \Delta^{k-1} g(j+1). \tag{2.11}$$

Now an analogue of Lemma 2 can be formulated.

Lemma 2.2. *Let* $\beta_1 > 0$, $\|f\|_1 < \infty$, *and*

$$\theta_s = \mu^{-1} \sum_{k=2}^s k|\beta_k| 2^{k-2} < 1/2. \tag{2.12}$$

Then the solution of the Stein equation

$$(\mathcal{A}_s g)(j) = f(j) - \pi_s\{f\}, \qquad j = 0, 1, \ldots, \tag{2.13}$$

has the following properties:

$$g(i) = 0, \quad i \leqslant 0; \tag{2.14}$$

$$\|\Delta g\|_1 \leqslant 2\|f\|_1(1-2\theta_s)^{-1}\mu^{-1}; \tag{2.15}$$

$$\|g\| \leqslant 2\|f\|_1(1-2\theta_s)^{-1}\mu^{-1}; \tag{2.16}$$

$$\|g\| \leqslant 2\|f\|(1-2\theta_s)^{-1}\mu^{-1/2}; \tag{2.17}$$

$$\|Ug\|_1 \leqslant 2\theta_s\|f\|_1(1-2\theta_s)^{-1}\mu^{-1}. \tag{2.18}$$

Possible applications of Lemma 2.2 are given in Section 4.

We first discuss how to replace recursively non-uniform estimates by uniform ones.

If $\|f\|_1 < \infty$, then from (2.13) it follows that in order to obtain a uniform local estimate, it suffices to estimate

$$|\,\mathrm{E}(\mathcal{A}_s g)(W)\,| = \left|\sum_{j=0}^{\infty} f(j)(P(j) - \pi_s(j))\right|.$$

Very slight changes are required for non-uniform estimates. Let $k \leqslant s$, and let

$$h_k(j) := (j-\mu)^k g(k), \qquad \varphi_k^{(s)}(j) := (\mathcal{A}_s g)(j)(j-\mu)^k - (\mathcal{A}_s h_k)(j). \tag{2.19}$$

Multiplying (2.13) by $(j-\mu)^k\pi_s(j)$ and summing over all non-negative j, we obtain

$$\sum_j (\mathcal{A}_s g)(j)(j-\mu)^k\pi_s(j) = \sum_j f(j)(j-\mu)^k\pi_s(j) - \pi_s\{f\}\mathrm{E}(W-\mu)^k. \tag{2.20}$$

Note that, due to the choice of $\beta_1, \ldots, \beta_s$, we have

$$\pi_s\{\mathcal{A}_s h_k\} = \sum_{j=0}^{\infty}(\mathcal{A}_s h_k)(j)\pi_s(j) = 0. \tag{2.21}$$

Thus, multiplying (2.13) by $(j-\mu)^k P(j)$, summing over all j, and subtracting the result obtained from (2.20) leads us to the inequality

$$\left|\sum_{j=0}^{\infty} f(j)(j-\mu)^k(P(j)-\pi_s(j))\right| \leqslant |\,\mathrm{E}(\mathcal{A}_s h_k)(W) + \left|\sum_{j=0}^{\infty}\varphi_k^{(s)}(j)(P(j)-\pi_s(j))\right|. \tag{2.22}$$

The maximum power of $(j-\mu)$ in $\varphi_k^{(s)}(j)$ is $k-1$. Therefore, the second summand in (2.22) can be estimated recursively, the first estimate being the estimate in total variation. Note that the non-uniform estimates in (2.22) are of the pseudo-moment type and correspond to the estimates considered by Barbour (1987).

In principle, the proofs of Lemmas 2.1 and 2.2 repeat that of Lemma 2.1 of Barbour and Xia (1999). Therefore, we prove Lemma 2.2 only.

Proof of Lemma 2.2. Let $\|f\|_1 < \infty$. If g_0 solves

$$\beta_1 g_0(j+1) - jg_0(j) = f(j) - \pi_1\{f\}, \qquad g_0(j) = 0,\ j < 0, \tag{2.23}$$

then

$$\|\Delta g_0\| \leqslant 2\|f\|_1\beta_1^{-1}; \tag{2.24}$$

see Barbour and Jensen (1989). For bounded $g : \mathbb{Z}_+ \to \mathbb{R}$, let Tg denote the solution $\tilde{g}$ to the equation

$$\beta_1 \tilde{g}(j+1) - j\tilde{g}(j) = f(j) - (Ug)(j) - \pi_1\{f\} + \pi_1\{Ug\}. \tag{2.25}$$

Note that $\|Ug_0\|_1 < \infty$ by (2.11), and we can recursively define $g_n = Tg_{n-1}$, $n \geq 0$, $g_{-1} = 0$. Set

$$\varkappa_n(j) := g_n(j) - g_{n-1}(j), \qquad j \geq 0.$$

Then

$$\beta_1 \varkappa_n(j+1) - j\varkappa_n(j) = -(U\varkappa_{n-1})(j) + \pi_1\{U\varkappa_{n-1}\}. \tag{2.26}$$

Applying (2.24), we obtain

$$\begin{aligned}\|\Delta\varkappa_n\|_1 &\leq 2\beta_1^{-1}\|U\varkappa_{n-1}\|_1 \leq 2\beta_1^{-1}\sum_{k=2}^{s} k|\beta_k|2^{k-2}\|\Delta\varkappa_{n-1}\|_1 \\ &= 2\theta_s\|\Delta\varkappa_{n-1}\|_1 \leq (2\theta_s)^n\|\Delta g_0\|_1 \leq 2\|f\|_1\beta_1^{-1}(2\theta_s)^n. \end{aligned} \tag{2.27}$$

Since $\theta_s < 1/2$, the limit $g_f = \lim_{n\to\infty} g_n$ exists and satisfies $Tg_f = g_f$. Furthermore,

$$\|\Delta g_f\|_1 \leq \sum_{n\geq 0}\|\Delta\varkappa_n\|_1 \leq \sum_{n\geq 0}(2\theta_s)^n 2\|f\|_1\beta_1^{-1} = 2\|f\|_1\beta_1^{-1}(1-2\theta_s)^{-1}. \tag{2.28}$$

Now $Tg_f = g_f$ is equivalent to

$$(\mathcal{A}_s g_f)(j) = f(j) - \pi_1\{f - Ug_f\}. \tag{2.29}$$

But

$$\pi_s\{\mathcal{A}_s g_f\} = 0 = \pi_s\{f\} - \pi_1\{f - Ug_f\}. \tag{2.30}$$

The proof of (2.15) follows from (2.25)–(2.30). The proof of (2.18) follows from (2.15) and the definition of U. The proof of (2.16) is evident from the inequality

$$|g(j)| = \left|\sum_{l=0}^{j-1}\Delta g(l)\right| \leq \|\Delta g\|_1.$$

Finally, (2.17) can be proved exactly as in Lemma 2.1 of Barbour and Xia (1999). □

3. Local estimates of the difference operator

In this section, we consider the local analogues of d_k. For W defined as in Section 1, let

$$P(j) := P(W = j), \qquad P_i(j) = P(W^{(i)} = j).$$

It is evident that

$$d_k = \|\Delta^k P\| = \sum_{j=0}^{\infty}|\Delta^k P(j)|. \tag{3.1}$$

Therefore, $|\Delta^k P(j)|$ is a local analogue of d_k. Obviously, any local estimate is majorized by the estimate in total variation

$$\sup_j |\Delta^k P(j)| \leqslant d_k. \tag{3.2}$$

Usually application of (3.2) reduces the accuracy of estimation. We shall show that $\sup_j |\Delta^k P(j)| = O(V^{-k/2-1/2})$ as $V \to \infty$, that is, it vanishes faster than d_k, which is of order $O(V^{-k/2})$. Moreover, we shall show that non-uniform estimates can be obtained as well as uniform ones. The estimates are formulated in terms of d_k and factorial moments of Z_i.

Set

$$a_s = 2\mu^{-1}\left\{ s d_{s-1} + d'_{s+1} \sum_{i=1}^{n} \left(\mu_i^2 + \mathrm{E}(Z_i)_{(2)}\right)\right\}, \tag{3.3}$$

$$b_s = 2d_s + 3\left\{ s a_{s-1} + \sum_{i=1}^{n} a_{s+1}^{(i)} \left(\mu_i^2 + \mathrm{E}(Z_i)_{(2)}\right)\right\}, \tag{3.4}$$

$$c_s = 6\mu a_s + 2d_s + 3s(a_{s-1} + b_{s-1}) + 1.5 \sum_{i=1}^{n} \Big\{ 2(a_{s+1}^{(i)}\mu_i + b_{s+1}^{(i)})(\mu_i^2 + \mathrm{E}(Z_i)_{(2)})$$
$$+\ a_{s+1}^{(i)}(\mu_i \mathrm{E}(Z_i+1)_{(2)} + \mathrm{E}Z_i^2(Z_i - 1))\Big\}. \tag{3.5}$$

The quantities $a_s^{(i)}$ and $b_s^{(i)}$ are defined by (3.3) and (3.4) with W replaced by $W^{(i)}$. For example,

$$a_s^{(i)} = 2(\mu - \mu_i)^{-1}\left\{ s d'_{s-1} + d''_{s+1} \sum_{j \neq i} \left(\mu_j^2 + \mathrm{E}(Z_j)_{(2)}\right)\right\}.$$

Theorem 3.1. *Let $\|f\|_1 < \infty$. The following estimates hold:*

(i) *If $\mathrm{E}Z_i^2 < \infty$, $i = 1, 2, \ldots, n$, and $s \geqslant 1$, then*

$$\left| \sum_{j=0}^{\infty} f(j) \Delta^s P(j-s) \right| \leqslant \|f\|_1 a_s. \tag{3.6}$$

(ii) *If $\mathrm{E}Z_i^2 < \infty$, $i = 1, 2, \ldots, n$, and $s \geqslant 2$, then*

$$\left| \sum_{j=0}^{\infty} f(j)(j-\mu) \Delta^s P(j-s) \right| \leqslant \|f\|_1 b_s, \tag{3.7}$$

$$\left| \sum_{j=0}^{\infty} f(j)(j-\mu) \Delta^s P(j-s) \right| \leqslant \|f\| d_s. \tag{3.8}$$

(iii) *If $\mathrm{E}Z_i^3 < \infty$, $i = 1, 2, \ldots, n$, and $s \geqslant 3$, then*

$$\left|\sum_{j=0}^{\infty} f(j)(j-\mu)^2\Delta^s P(j-s)\right| \leqslant \|f\|_1 c_s. \tag{3.9}$$

In particular, when $f(j)$ is the indicator of a point, we obtain the following corollary.

Corollary 3.1. *If* $\mathrm{E}Z_i^2 < \infty$, $i = 1, \ldots, n$, *and* $s \geqslant 2$, *then*

$$\sup_{j\geqslant 0} |\,\Delta^s P(j)\,| \leqslant a_s, \tag{3.10}$$

$$\sup_{j\geqslant 0} |\,(j-\mu)\Delta^s P(j-s)\,| \leqslant b_s. \tag{3.11}$$

Remark 3.1. Suppose that Z_i do not depend on n, $\mathrm{E}Z_i^3 < \infty$, and $v_i > 0$ $(i = 1, 2, \ldots, n)$. Then

$$a_s = O(n^{-(s+1)/2}), \qquad b_s = O(n^{-s/2}), \qquad c_s = O(n^{-(s-1)/2}).$$

Remark 3.2. In principle, applying the proof of Theorem 3.1 recursively, we can obtain estimates for all

$$\left|\sum_j f(j)(j-\mu)^k \Delta^s P(j-s)\right|, \qquad k \leqslant s.$$

However, with the growth of k, the estimates become more and more complicated and, as can be seen from Corollary 3.1, the order of accuracy decreases.

Proof of Theorem 3.1. All estimates are obtained similarly. Therefore, we prove only (3.9). We begin with the Stein equation for the Poisson approximation

$$(\mathcal{A}_1 g)(j) = \mu g(j+1) - jg(j) = f(j) - \pi_1\{f\}. \tag{3.12}$$

Let us multiply (3.12) by $(j-\mu)^2\Delta^s P(j-s)$ and sum over j. Note that $P(j-m) = 0$ for $j < m$. Consequently,

$$\begin{aligned}\sum_{j=0}^{\infty}(j-\mu)^2 P(j-\mu) &= \sum_{j=m}^{\infty}(j-\mu)^2 P(j-\mu)\\ &= \sum_{k=0}^{\infty}(k-\mu)^2 P(k) + 2m\left(\sum_{k=0}^{\infty} kP(k) - \mu\right) + m^2 = \sigma^2 + m^2\end{aligned}$$

and

$$\sum_{j=0}^{\infty}(j-\mu)^2\Delta^s P(j-s) = \sum_{j=0}^{\infty}(j-\mu)^2\sum_{m=0}^{s}\binom{s}{m}(-1)^m P(j-m)$$
$$= \sum_{m=0}^{s}\binom{s}{m}(-1)^m\sum_{j=0}^{\infty}(j-\mu)^2 P(j-m)$$
$$= \sum_{m=0}^{s}\binom{m}{j}(-1)^m(\sigma^2+m^2) = 0.$$

For the proof of the last equality, note that $s \geqslant 3$ and, consequently,

$$\sum_{m=0}^{s}\binom{s}{m}(-1)^m m^2 = \sum_{m=0}^{s}\binom{s}{m}(-1)^m m(m-1)$$
$$= s(s-1)\sum_{m=2}^{s}\binom{s-2}{m-2}(-1)^{m-2} = 0.$$

Moreover,

$$(j-\mu)g(j) = \mu\Delta g(j) - f(j) + \pi_1\{f\}, \tag{3.13}$$
$$\Delta^k(\mathcal{A}_1 h_2)(j) = (\mathcal{A}_1\Delta^k h_2)(j) - k\Delta^{k-1}h_2(j+1).$$

Therefore,

$$\left|\sum_{j=0}^{\infty} f(j)(j-\mu)^2\Delta^s P(j-s)\right| = \left|\sum_{j=0}^{\infty}(\mathcal{A}_1 g)(j)(j-\mu)^2\Delta^s P(j-s)\right|$$
$$\leqslant \left|\sum_{j=0}^{\infty}(\mathcal{A}_1 h_2)(j)\Delta^s P(j-s)\right| + \left|\sum_{j=0}^{\infty}\varphi_2^{(1)}(j)\Delta^s P(j-s)\right|$$
$$= \left|\sum_{j=0}^{\infty} P(j)\Delta^s(\mathcal{A}_1 h_2)(j)\right| + \left|\sum_{j=0}^{\infty}\varphi_2^{(1)}(j)\Delta^s P(j-s)\right|$$
$$\leqslant |\,\mathrm{E}(\mathcal{A}_1\Delta^s h_2)(W)\,| + s|\,\mathrm{E}\Delta^{s-1}h_2(W+1)\,|$$
$$+ \left|\sum_{j=0}^{\infty}\varphi_2^{(1)}(j)\Delta^s P(j-s)\right|. \tag{3.14}$$

Here

$$\varphi_2^{(1)}(j) = (j-\mu)^2(\mathcal{A}_1 g)(j) - (\mathcal{A}_1 h_2)(j) = -\mu g(j+1)(2(j+1)-2\mu-1)$$
$$= 2\mu f(j+1) - 2\mu\pi_1\{f\} - 2\mu^2\Delta g(j+2) - \mu g(j+1).$$

Note that, for π_1, (2.15)–(2.16) hold with $\theta_1 = 0$. Consequently,

$$\left|\sum_{j=0}^{\infty} \varphi_2^{(1)}(j)\Delta^s P(j-s)\right| \leqslant 2\mu(\|f\|_1 + \mu\|\Delta g\|_1)\sup_j |\Delta^s P(j-s)|$$

$$+ \mu\|g\| \sum_{j=0}^{\infty} |\Delta^s P(j-s)| \leqslant \|f\|_1(6a_s + 2d_s). \tag{3.15}$$

Similarly,

$$|\mathrm{E}\Delta^{s-1} h_2(W+1)| = \left|\sum_{j=0}^{\infty}(j-\mu)^2 g(j)\Delta^{s-1} P(j-s)\right|$$

$$\leqslant (\|f\|_1 + \mu\|\Delta g\|_1)\sup_j |(j-\mu)\Delta^{s-1} P(j-s)|$$

$$\leqslant 3\|f\|_1(b_{s-1} + a_{s-1}). \tag{3.16}$$

Finally,

$$h_2(w+l) = h_2(w+1) + \sum_{m=1}^{l-1} \Delta h_2(w+m)$$

and

$$\mathrm{E}h_2(W+1) = \sum_{j=0}^{\infty} q_{ij}\mathrm{E}h_2(W^{(i)} + j + 1)$$

$$= \mathrm{E}h_2(W^{(i)} + 1) + \sum_{j=0}^{\infty} q_{ij} \sum_{m=1}^{j} \mathrm{E}\Delta h_2(W^{(i)} + m),$$

$$\mathrm{E}Z_i h_2(W) = \mu_i \mathrm{E}h_2(W^{(i)} + 1) + \sum_{j=0}^{\infty} q_{ij} j \sum_{m=1}^{j-1} \mathrm{E}\Delta h_2(W^{(i)} + m).$$

Therefore,

$$|\mathrm{E}(\mathcal{A}_1\Delta^s h_2)(W)| \leqslant \sum_{\mathrm{i}=1}^{\mathrm{n}} |\mu_\mathrm{i}\mathrm{E}(\Delta^\mathrm{s}\mathrm{h}_2)(\mathrm{W}+1) - \mathrm{E}\mathrm{Z}_\mathrm{i}(\Delta^\mathrm{s}\mathrm{h}_2)(\mathrm{W})|$$

$$\leqslant \sum_{i=1}^{n}\sum_{j=0}^{\infty} q_{ij}\left(\mu_i \sum_{m=1}^{j} |\mathrm{E}\Delta^{s+1} h_2(W^{(i)} + m)| + j\sum_{m=1}^{j-1} |\mathrm{E}\Delta^{s+1} h_2(W^{(i)} + m)|\right). \tag{3.17}$$

By (3.13) we have

$$|\,\mathrm{E}\Delta^{s+1}h_2(W^{(i)}+m)\,| = \left|\sum_{j=0}^{\infty}(j-\mu)^2 g(j)\Delta^{s+1}P_i(j-s-1-m)\right|$$
$$\leqslant (\|f\|_1+\mu\|\Delta g\|_1)\sup_j|(j-\mu)\Delta^{s+1}P_i(j-s-1-m)|$$
$$\leqslant 3\|f\|_1(b_{s+1}^{(i)}+(m+\mu_i)a_{s+1}^{(i)}). \tag{3.18}$$

Estimate (3.9) follows from (3.14)–(3.18). □

4. Approximation by π_s

Let π_s be defined as in Section 2. Recall that $\beta_j=\sum_{i=1}^{n}\beta_j^{(i)}$, where the jth factorial cumulant of Z_i equals $\beta_j^{(i)}j!$. The quantities $a_s^{(i)}$ and $b_s^{(i)}$ are the same as in Theorem 3.1. Set

$$\varepsilon_1^{(i)}=\mathrm{E}(Z_i)_{(s+1)}/s!+\sum_{k=1}^{s}k|\beta_k^{(i)}|\,\mathrm{E}(Z_i)_{(s-k+1)}/(s-k+1)!,$$

$$\varepsilon_2^{(i)}=\mathrm{E}Z_i(Z_i)_{(s+1)}/(s+1)!+\sum_{k=1}^{s}k|\beta_k^{(i)}|\,\mathrm{E}(Z_i+1)_{(s-k+2)}/(s-k+2)!,$$

$$K_1=\sum_{k=1}^{s}k(k+1)|\,\beta_k\,|2^k,$$

$$K_2=3\sum_{k=1}^{s}k(k+1)|\,\beta_k\,|2^k+\mu^{-1/2}\sum_{k=1}^{s}k^2(k+3)|\beta_k|^{k-1}.$$

Now we can formulate the main result of this paper.

Theorem 4.1. *Let $\|f\|_1<\infty$, $s\geqslant 3$, $\mu>0$, $\theta_s<1/2$, and $\mathrm{E}Z_i^{s+1}<\infty$, $i=1,\ldots,n$. Then the following estimates hold:*

$$\|\mathcal{L}(W)-\pi_s\|\leqslant\frac{2\mathrm{d}_\mathrm{s}'}{1-2\theta_\mathrm{s}}\mu^{-1/2}\sum_{i=1}^{n}\varepsilon_1^{(i)}; \tag{4.1}$$

$$\left|\sum_{j=0}^{\infty}f(j)(P(j)-\pi_s(j))\right|\leqslant 2\|f\|_1\frac{d_s'}{1-2\theta_s}\mu^{-1}\sum_{i=1}^{n}\varepsilon_1^{(i)}; \tag{4.2}$$

$$\left|\sum_{j=0}^{\infty}f(j)(j-\mu)(P(j)-\pi_s(j))\right|\leqslant\frac{\|f\|_1}{1-2\theta_s}\left(\frac{d_s'}{1-2\theta_s}K_1\mu^{-3/2}\sum_{i=1}^{n}\varepsilon_1^{(i)}+3\sum_{i=1}^{n}a_s^{(i)}\varepsilon_1^{(i)}\right); \tag{4.3}$$

$$\left|\sum_{j=0}^{\infty} f(j)(j-\mu)^2(P(j)-\pi_s(j))\right| \leqslant \|f\|_1 \frac{d_s'}{(1-2\theta_s)^2} K_2\mu^{-1}\sum_{i=1}^{n}\varepsilon_1^{(i)}$$

$$+3\frac{\|f\|_1}{1-2\theta_s}\sum_{i=1}^{n}\left\{(a_s^{(i)}\mu_i+b_s^{(i)})\varepsilon_1^{(i)}+a_s^{(i)}\varepsilon_2^{(i)}\right\}. \qquad (4.4)$$

The accuracy of approximation is most evident in the case of independent Bernoulli variables. In this case, $\beta_j = (-1)^{j+1}\sum_{i=1}^n p_i^j/j$.

Corollary 4.1. *Let* $Z_i = I_i$, $P(I_i = 1) = p_i \leqslant 1/5(i = 1, \ldots, n)$. *Let* $s \geqslant 3$, *and let* $\beta_1 = \sum_{i=1}^n p_i \to \infty$ *as* $n \to \infty$. *Then*

$$\|\mathcal{L}(W)-\pi_s\| = O\left(\beta_1^{-(s+1)/2}\sum_{i=1}^{n} p_i^{s+1}\right),$$

$$\sup_{j\geqslant 0}|P(j)-\pi_s(j)| = O\left(\beta_1^{-(s+2)/2}\sum_{i=1}^{n} p_i^{s+1}\right),$$

$$\sup_{j\geqslant 0}|(j-\beta_1)(P(j)-\pi_s(j))| = O\left(\beta_1^{-(s+2)/2}\sum_{i=1}^{n} p_i^{s+1}\right),$$

$$\sup_{j\geqslant 0}|(j-\beta_1)^2(P(j)-\pi_s(j))| = O\left(\beta_1^{-s/2}\sum_{i=1}^{n} p_i^{s+1}\right).$$

Corollary 4.1 shows that the SCP approximation π_s is quite accurate even for $p_i = O(1)$, that is, in the case where the standard Poisson approximation fails. Note that we do not investigate the convergence to the Poisson law, that is, the case of small p_i.

Finally, a few words should be said about the condition $\theta_s < 1/2$. In principle, this condition is quite restrictive. For $s = 2$, it implies that $\mu < 2\sigma^2 < 3\mu$. On the other hand, if the factorial cumulant of Z_i (recall that this equals $\beta_j^{(i)} j!$), for some sufficiently large A, satisfies the standard condition of large deviations

$$|\beta_j^{(i)}| \leqslant \frac{\mu_i}{jA^{j-1}}, \qquad (4.5)$$

then $\theta_s < 1/2$. Note that the Bernoulli variable I_i satisfies (4.5) with $A = 1/(\max p_i)$.

For the proof of Theorem 4.1 we need an auxiliary lemma.

Lemma 4.1. *Let the assumptions of Theorem* 4.1 *be satisfied. Let* h *be either of the functions* h_1 *or* h_2. *Then*

$$|\,\mathrm{E}(\mathcal{A}_s h)(W)\,| \leqslant \sum_{i=1}^{n}\sum_{j=0}^{\infty} q_{ij}\Bigg\{\sum_{k=1}^{s} k|\beta_k^{(i)}|\sum_{m=1}^{j-s+k}\binom{j-m}{s-k}|\,\mathrm{E}\Delta^s h(W^{(i)}+m)\,|$$
$$+\, j\sum_{m=1}^{j-s}\binom{j-m-1}{s-1}|\,\mathrm{E}\Delta^s h(W^{(i)}+m)\,|\Bigg\}. \qquad (4.6)$$

Proof. As usual, we assume that $\sum_a^b = 0$, $b < a$. By (2.10) we have

$$|\,\mathrm{E}(\mathcal{A}_s h)(W)\,| \leqslant \sum_{i=1}^{n}\left|\sum_{k=1}^{s} k\beta_k^{(i)}\mathrm{E}\Delta^{k-1}h(W+1) - \mathrm{E}Z_i h(W)\right|. \qquad (4.7)$$

By Newton's formula

$$h(w+l) = h(w+1) + \sum_{m=1}^{t}\binom{l-1}{m}\Delta^m h(w+1) + \sum_{m=1}^{l-t-1}\binom{l-1-m}{t}\Delta^{t+1}h(w+m),$$

we obtain

$$\mathrm{E}\Delta^{k-1}h(W+1) = \sum_{j=0}^{\infty} q_{ij}\mathrm{E}\Delta^{k-1}h(W^{(i)}+j+1)$$
$$= \sum_{l=k-1}^{s-1}\mathrm{E}\Delta^l h(W^{(i)}+1)\mathrm{E}(Z_i)_{(l+1-k)}/(l+1-k)!$$
$$+ \sum_{j=0}^{\infty} q_{ij}\sum_{m=1}^{j-s+k}\binom{j-m}{s-k}\mathrm{E}\Delta^s h(W^{(i)}+m) \qquad (4.8)$$

and

$$\mathrm{E}Z_i h(W) = \sum_{j=0}^{\infty} jq_{ij}\mathrm{E}h(W^{(i)}+j)$$
$$= \sum_{m=1}^{s-1}\mathrm{E}\Delta^m h(W^{(i)}+1)\mathrm{E}(Z_i)_{(m+1)}/m!$$
$$+ \sum_{j=0}^{\infty} jq_{ij}\sum_{m=1}^{j-s}\binom{j-m-1}{s-1}\mathrm{E}\Delta^s h(W^{(i)}+m). \qquad (4.9)$$

Changing the order of summation, we obtain

$$\sum_{k=1}^{s} k\beta_k^{(i)} \sum_{l=k-1}^{s-1} \mathrm{E}\Delta^l h(W^{(i)}+1)\mathrm{E}(Z_i)_{(l+1-k)}/(l+1-k)!$$

$$= \sum_{m=0}^{s-1} \mathrm{E}\Delta^m h(W^{(i)}+1) \sum_{k=1}^{m+1} k\beta_k^{(i)}\mathrm{E}(Z_i)_{(m+1-k)}/(m+1-k)!. \tag{4.10}$$

However, the identity

$$\sum_{k=1}^{m+1} k\beta_k^{(i)}\mathrm{E}(Z_i)_{(m+1-k)}/(m+1-k)! = \mathrm{E}(Z_i)_{(m+1)}/m!. \tag{4.11}$$

holds, and this can be proved as follows. Denote by $\psi(z)$ the generating function of Z_i. Expanding $\psi(z)$ in factorial moments and $\ln\psi(z)$ in factorial cumulants, comparing the coefficients to $(z-1)^k$, and applying the identity $\psi'(z) = \psi(z)(\ln\psi(z))'$, we obtain (4.11). Estimate (4.6) follows from (4.8)–(4.11). □

Proof of Theorem 4.1. We shall prove only (4.4). Set

$$M(j) = \mu\Delta g(j+1) + U(j) - f(j). \tag{4.12}$$

Then by (3.1) we have

$$\|M\|_1 \leqslant 3\|f\|_1(1-2\theta_s)^{-1}. \tag{4.13}$$

We shall apply (2.3) and (2.6), with λ_j defined by (2.9). As above, $h_2(j) = (j-\mu)^2 g(j)$. Set

$$\varphi_2(j) = (\mathcal{A}g)(j)(j-\mu)^2 - (\mathcal{A}h_2)(j).$$

Applying (4.13), we obtain

$$\varphi_2(j) = -\sum_{l=1}^{s} l^2\lambda_l g(j+l)(2j-2\mu+l)$$

$$= -2\sum_{l=1}^{s} l^2\lambda_l(M(j+l)+\pi_s\{f\}) + \sum_{l=1}^{s} l^3\lambda_l g(j+l). \tag{4.14}$$

Now we have

$$\left|\sum_{j=0}^{\infty} f(j)(j-\mu)^2(P(j)-\pi_s(j))\right| \leqslant |\mathrm{E}(\mathcal{A}_s h_2)(W)| + \left|\sum_{j=0}^{\infty} \varphi_2(j)(P(j)-\pi_s(j))\right|. \tag{4.15}$$

Taking into account (4.14) and (2.16), we see that the second summand in the right-hand side of (4.15) is majorized by

$$2\|M\|_1 \sum_{l=1}^{s} l^2|\lambda_l| \sup_j |P(j) - \pi_s(j)| + \|g\| \|\mathcal{L}(W) - \pi_s\| \sum_{l=1}^{s} l^3|\lambda_l|$$

$$\leqslant \|f\|_1 d'_s \mu^{-1}(1-2\theta_s)^{-2} K_2 \sum_{i=1}^{n} \varepsilon_1^{(i)}. \tag{4.16}$$

In (4.16), we used the estimate

$$\sum_{l=1}^{s} l^2|\lambda_l| \leqslant \sum_{l=1}^{s} l^2 \sum_{k=1}^{s} \binom{k}{l} |\beta_k| = \sum_{k=1}^{s} k(k+1)2^{k-2}|\beta_k|$$

and a similar estimate for the second summand.

Similarly, we have

$$\begin{aligned} |\mathrm{E}\Delta^s h_2(W^{(i)} + m)| &= \left| \sum_{j=0}^{\infty} P_i(j) \Delta^s h_2(j+m) \right| \\ &= \left| \sum_{j=0}^{\infty} (j+m-\mu)^2 g(j+m) \Delta^s P_i(j-s) \right| \\ &= \left| \sum_{j=0}^{\infty} (j+m-\mu) M(j+m) \Delta^s P_i(j-s) \right| \\ &\leqslant \|M\|_1 \sup_j |(j+m-\mu)\Delta^s P_i(j-s)| \\ &\leqslant 3\|f\|_1 (1-2\theta_s)^{-1}(b_s^{(i)} + (m+\mu_i)a_s^{(i)}). \end{aligned} \tag{4.17}$$

To finish the proof, it suffices to use Lemma 4.1 and substitute the estimate obtained and (4.16) into (4.15). □

5. Approximation of dependent indicators

We shall formulate one non-uniform result for a Poisson approximation of possibly dependent indicators. Note that, in this case, the approach of Barbour *et al*. (1995) is inapplicable. For the indicator variable, we use the notation I_i as above. However, emphasizing the possible dependence of indicators, we denote their sum by $\tilde{W}$, that is,

$$\tilde{W} = \sum_{i=1}^{n} I_i, \qquad P(I_i = 1) = 1 - P(I_i = 0) = p_i < 1,$$

$$\tilde{W}^{(i)} = \tilde{W} - I_i, \qquad \mu = \sum_{i=1}^{n} p_i.$$

As before, π_1 denotes the standard Poisson approximation with parameter μ. Let V_i be constructed on the same probability space as $(\tilde{W})^{(i)}$ so that

$$P(V_i = j) = P(\tilde{W}^{(i)} | I_i = j).$$

Proposition 5.1. *For any $j \in \mathbb{Z}_+$,*

$$|j - \mu||P(\tilde{W} = j) - \pi_1(j)| \leqslant 2\|\mathcal{L}(\tilde{W}) - \pi_1\| + 6\sum_{i=1}^{n} p_i^2 \max_j |P(V_i = j)$$
$$- P(V_i = j-1)| + 2\sum_{i=1}^{n} p_i \max_j |P(\tilde{W}^{(i)} = j) - P(V_i = j)|. \tag{5.1}$$

Proposition (5.1) demonstrates which components are needed for a non-uniform estimate. The estimate in total variation can be found in Barbour *et al*. (1992b, Theorem 1B). However, for other summands, an analogue of Theorem 3.1 is required. In fact, obtaining an analogue of Theorem 3.1 is the hardest part in the dependent case. Therefore, Proposition 5.1 gives an insight to the problem rather than its solution.

Proof. We apply (2.22). By (2.16) we have

$$\left|\sum_{j=0}^{\infty} \varphi_1^{(1)}(j)(P(\tilde{W} = j) - \pi_1(j))\right| = \mu\left|\sum_{j=0}^{\infty} g(j+1)(P(\tilde{W} = j) - \pi_1(j))\right|$$
$$\leqslant \mu\|g\|\|\mathcal{L}(\tilde{W}) - \pi_1\| \leqslant 2\|\mathcal{L}(\tilde{W}) - \pi_1\|.$$

Furthermore, we have

$$|\mathrm{E}(\mathcal{A}_1 h_1)(\tilde{W})| = \left|\sum_{i=1}^{n} p_i\big(\mathrm{E}h_1(\tilde{W}+1) - \mathrm{E}\{h_1(\tilde{W}^{(i)}+1)|I_i = 1\}\big)\right|$$
$$= \left|\sum_{i=1}^{n} p_i^2 \mathrm{E}\{\Delta h_1(\tilde{W}^{(i)}+1)|I_i = 1\} + \sum_{i=1}^{n} p_i q_i (\mathrm{E}\{h_1(\tilde{W}^{(i)}+1)|I_i = 0\}.\right.$$
$$- \mathrm{E}\{h_1(\tilde{W}^{(i)}+1)|I_i = 1\})| \leqslant \sum_{i=1}^{n} p_i^2 \left|\sum_{j=0}^{\infty} \Delta h_1(j+1) P(V_i = j-1)\right|$$
$$+ \sum_{i=1}^{n} p_i \left|\sum_{j=0}^{\infty} h_1(j+1)(P(\tilde{W}^{(i)} = j) - P(V_i = j))\right| = J_1 + J_2.$$

Applying (3.13), we can estimate J_1 as follows:

$$J_1 = \sum_{i=1}^{n} p_i^2 |\, h(j)(P(V_i = j-2) - P(V_i = j-3))|$$

$$\times \sum_{i=1}^{n} p_i \left| \sum_{j=0}^{\infty} (f(j) - \pi_1\{f\} - \mu\Delta g(j))(P(V_i = j-2) - P(V_i = j-3)) \right|$$

$$\leqslant 2(1 + \mu \|\Delta g\|_1) \sum_{i=1}^{n} p_i^2 \max_j |P(V_i = j) - P(V_i = j-1)|$$

$$\leqslant 6 \sum_{i=1}^{n} p_i^2 \max_j |P(V_i = j) - P(V_i = j-1)|.$$

Similarly,

$$J_2 \leqslant 2 \sum_{i=1}^{n} p_i \max_j |P(\tilde{W}^{(i)} = j) - P(V_i = j)\,|.$$

□

Acknowledgement

I would like to thank the referee for some useful remarks.

References

Barbour, A.D. (1987) Asymptotic expansions in the Poisson limit theorem. *Ann. Probab.*, **16**, 748–766.

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

Barbour, A.D. and Chryssaphinou, O. (2001) Compound Poisson approximation: a user's guide. *Ann. Appl. Probab.*, **11**, 964–1002.

Barbour, A.D. and Jensen, J.L. (1989) Local and tail approximations near the Poisson limit. *Scand. J. Statist.*, **16**, 75–87.

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

Barbour, A.D., Chen, L.H.Y. and Choi, K.P. (1995) Poisson approximation for unbounded functions I: Independent summands. *Statist. Sinica*, **5**, 749–766.

Barbour, A.D., Chen, L.H.Y. and Loh, W. (1992a) Compound Poisson approximation for nonnegative random variables using Stein's method. *Ann. Probab.*, **20**, 1843–1866.

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

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

Chen, L.H.Y. (1975) Poisson approximation for dependent trials. *Ann. Probab.*, **3**, 534–545.

Chen, L.H.Y. (2000) Non-uniform bounds in probability approximations using Stein's method. Probability and Statistical Models with Applications: A Volume in Honour of Theophilos

Cacoullos, In: Ch.A. Charalambides, M.V. Koutras and N.Balakrishnan (eds), pp. 3–14. Boca Raton, FL: Chapman & Hall/CRC.

Chen, L.H.Y. and Choi, K.P. (1992) Some asymptotic and large deviations results in Poisson approximation. *Ann. Probab.*, **20**, 1867–1876.

Chen, L.H.Y. and Roos, M. (1995) Compound Poisson approximation for unbounded functions on a group, with application to large deviations. *Probab. Theory Related Fields*, **103**, 515–528.

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

Roos, B. (2002) Kerstan's method in the multivariate Poisson approximation: an expansion in the exponent. *Teor Veroyatnost. i Primen.*, **47**(2), 397–401.

Received April 2003 and revised November 2003