Zero bias transformation and asymptotic expansions II : the Poisson case

Ying Jiao

Zero bias transformation and asymptotic expansions II : the Poisson case

Ying Jiao*

April 27, 2009

Abstract

We apply a discrete version of the methodology in [12] to obtain a recursive asymptotic expansion for $\mathbb{E}[h(W)]$ in terms of Poisson expectations, where W is a sum of independent integer-valued random variables and h is a polynomially growing function. We also discuss the remainder estimations.

MSC 2000 subject classifications: 60G50, 60F05.
Key words: Poisson approximation, zero bias transformation, Stein's method, asymptotic expansions, discrete reverse Taylor formula.

1 Introduction and main result

It should be noted in the first place that the notation used in this paper is similar as in [12], however, their meanings are different since we here consider discrete random variables. Stein's method for Poisson appoximation has been introduced by Chen [8]. Let Z be an $\mathbb{N}$-valued random variable ($\mathbb{N}$-r.v.), then Z follows the Poission distribution with parameter λ if and only if the equality $\mathbb{E}[Zf(Z)] = \lambda\mathbb{E}[f(Z+1)]$ holds for any function $f : \mathbb{N} \to \mathbb{R}$ such that both sides of the equality are well defined. Based on this observation, Chen has proposed the following discrete Stein's equation:

$$xf(x) - \lambda f(x+1) = h(x) - \mathcal{P}_\lambda(h), \quad x \in \mathbb{N} \tag{1}$$

where $\mathcal{P}_\lambda(h)$ is the expectation of h with respect to the λ-Poisson distribution. If X is an $\mathbb{N}$-r.v., one has $\mathbb{E}[h(X)] - \mathcal{P}_\lambda(h) = \mathbb{E}[Xf_h(X) - f_h(X+1)]$ where f_h is a solution of (1) and is given as

$$f_h(x) = \frac{(x-1)!}{\lambda^x}\sum_{i=x}^{\infty}\frac{\lambda^i}{i!}\big(h(i) - \mathcal{P}_\lambda(h)\big). \tag{2}$$

*Laboratoire de probabilités et modèles aléatoires, Université Paris 7; jiao@math.jussieu.fr

The value $f_h(0)$ can be arbitrary and is not used in calculations in general.

Stein's method has been adopted for Poisson approximation problems since [8] in a series of papers such as [1], [5], [4] among many others, one can also consult the monograph [6] and the survey paper [10]. In particular, Barbour [3] has developed, in parallel with the normal case [2], asymptotic expansions for sum of independent $\mathbb{N}$-r.v.s and for polynomially growing functions. The asymptotic expansion problem has also been studied by using other methods such as Lindeberg method (e.g. [7]).

In this paper, we address this problem by the zero bias transformation approach. Similar as in Goldstein and Reinert [11], we introduce a discrete analogue of zero bias transformation (see also [9]). Let X be an $\mathbb{N}$-r.v. with expectation λ. We say that an $\mathbb{N}$-r.v. X^* has Poisson X-zero biased distribution if the equality

$$\mathbb{E}[Xf(X)] = \lambda \mathbb{E}[f(X^* + 1)] \tag{3}$$

holds for any function $f : \mathbb{N} \to \mathbb{R}$ such that the left side of (3) is well defined. The distribution of X^* is unique: one has $\mathbb{P}(X^* = x) = (x+1)\mathbb{P}(X = x+1)/\lambda$. Combining Stein's equation (1) and zero bias transformation (3), the error of the Poisson approximation can be written as

$$\mathbb{E}[h(X)] - \mathcal{P}_\lambda(h) = \lambda \mathbb{E}[f_h(X^* + 1) - f_h(X + 1)]. \tag{4}$$

A first order correction term for the Poisson approximation has been proposed in [9] by using the Poisson zero bias transformation.

Recall the difference operator Δ defined as $\Delta f(x) = f(x+1) - f(x)$. For any $x \in \mathbb{N}_* := \mathbb{N} \setminus \{0\}$ and any $n \in \mathbb{N}$, one has $\Delta\binom{x}{n} = \binom{x}{n-1}$. If f and g are two functions on $\mathbb{N}$, then

$$\Delta(f(x)g(x)) = f(x+1)\Delta g(x) + g(x)\Delta f(x).$$

We have the Newton's expansion ([4, Thm5.1]), which can be viewed as an analogue of the Taylor's expansion in the discrete case. For all $x, y \in \mathbb{N}$ and $N \in \mathbb{N}$,

$$f(x+y) = \sum_{j=0}^{N} \binom{y}{j} \Delta^j f(x) + \sum_{0 \le j_1 < \cdots < j_{N+1} < y} \Delta^{N+1} f(x + j_1).$$

Let us introduce the following quantity, where we use the same notation as in [12], but its meaning is changed. For any $\mathbb{N}$-r.v. Y and any $k \in \mathbb{N}$ such that $\mathbb{E}[|Y|^k] < +\infty$, denote by

$$m_Y^{(k)} := \mathbb{E}\Big[\binom{Y}{k}\Big] = k!\,[Y]_k \tag{5}$$

where $[Y]_k$ is the k^{th} factorial moment of Y. Let X and Y be two independent $\mathbb{N}$-r.v.s and $f : \mathbb{N} \to \mathbb{R}$ such that $\Delta^k f(X)$ and $\Delta^k f(X+Y)$ are both integrable, then

$$\mathbb{E}[f(X+Y)] = \sum_{k=0}^{N} m_Y^{(k)} \mathbb{E}[\Delta^k f(X)] + \delta_N(f, X, Y). \tag{6}$$

where

$$\delta_N(f,X,Y)=\mathbb{E}\Big[\sum_{0\le j_1<\cdots<j_{N+1}<Y}\Delta^{N+1}f(X+j_1)\Big]. \tag{7}$$

We introduce the discrete reverse Taylor formula. Once again, the following result is very similar with [12, Pro1.1], however, with different significations of notation.

Proposition 1.1 (discrete reverse Taylor formula) *With the above notation, we have*

$$\mathbb{E}[f(X)]=\sum_{d\ge 0}(-1)^d\sum_{\mathbf{J}\in\mathbb{N}_*^d,\,|\mathbf{J}|\le N}m_Y^{(\mathbf{J})}\mathbb{E}[\Delta^{|\mathbf{J}|}f(X+Y)]+\varepsilon_N(f,X,Y) \tag{8}$$

where

$$\varepsilon_N(f,X,Y):=-\sum_{d\ge 0}(-1)^d\sum_{\mathbf{J}\in\mathbb{N}_*^d,\,|\mathbf{J}|\le N}m_Y^{(\mathbf{J})}\delta_{N-|\mathbf{J}|}(\Delta^{|\mathbf{J}|}f,X,Y), \tag{9}$$

for any integer $d\ge 1$ and any $\mathbf{J}=(j_l)_{l=1}^d\in\mathbb{N}_*^d$, $|\mathbf{J}|=j_1+\cdots+j_d$ and $m_Y^{(|\mathbf{J}|)}:=m_Y^{(j_1)}\cdots m_Y^{(j_d)}$, and by convention, $\mathbb{N}_*^0=\{\emptyset\}$ with $|\emptyset|=0$, $m_Y^{(\emptyset)}=1$.

Consider now a family of independent $\mathbb{N}$-r.v.s $X_i\,(i=1,\cdots n)$ with expectations λ_i, which are "sufficiently good" in a sense we shall precise later. Let $W=X_1+\cdots+X_n$ and denote $\lambda_W:=\mathbb{E}[W]=\lambda_1+\cdots+\lambda_n$. Let $W^{(i)}=W-X_i$ and X_i^* be an $\mathbb{N}$-r.v., independent of $W^{(i)}$ and which has the Poisson X_i-zero biased distribution. Finally, let I be a random index valued in $\{1,\cdots,n\}$ which is independent of $(X_1,\cdots,X_n,X_1^*,\cdots,X_n^*)$ and such that $\mathbb{P}(I=i)=\lambda_i/\lambda_W$ for any i. Then, similar as in [11], the random variable $W^*:=W^{(I)}+X_I^*$ follows the Poisson W-zero biased distribution.

We give below the asymptotic expansion formula in the Poisson case.

Theorem 1.2 *Let $N\in\mathbb{N}$ and $p\ge 0$. Let $h:\mathbb{N}\to\mathbb{R}$ be a function which is of $O(x^p)$ at infinity and $X_i\,(i=1,\cdots,n)$ be a family of independent $\mathbb{N}$-r.v.s having up to $(N+p+1)^{th}$ order moments. Let $W=X_1+\cdots+X_n$ and $\lambda_W=\mathbb{E}[W]$. Then $\mathbb{E}[h(W)]$ can be written as the sum of two terms $C_N(h)$ and $e_N(h)$ such that $C_0(h)=\mathcal{P}_{\lambda_W}(h)$ and $e_0(h)=\mathbb{E}[h(W)]-\mathcal{P}_{\lambda_W}(h)$, and recursively for any $N\ge 1$,*

$$C_N(h)=C_0(h)+\sum_{i=1}^n\lambda_i\sum_{d\ge 1}(-1)^{d-1}\sum_{\mathbf{J}\in\mathbb{N}_*^d,\,|\mathbf{J}|\le N}m_{X_i}^{(\mathbf{J}^\circ)}\big(m_{X_i^*}^{(\mathbf{J}^\dagger)}-m_{X_i}^{(\mathbf{J}^\dagger)}\big)C_{N-|\mathbf{J}|}(\Delta^{|\mathbf{J}|}f_h(x+1)), \tag{10}$$

$$\begin{aligned}e_N(h)=\sum_{i=1}^n\lambda_i\Big[&\sum_{d\ge 1}(-1)^{d-1}\sum_{\mathbf{J}\in\mathbb{N}_*^d,\,|\mathbf{J}|\le N}m_{X_i}^{(\mathbf{J}^\circ)}\big(m_{X_i^*}^{(\mathbf{J}^\dagger)}-m_{X_i}^{(\mathbf{J}^\dagger)}\big)e_{N-|\mathbf{J}|}(\Delta^{|\mathbf{J}|}f_h(x+1))\\&+\sum_{k=0}^N m_{X_i^*}^{(k)}\varepsilon_{N-k}(\Delta^k f_h(x+1),W^{(i)},X_i)+\delta_N(f_h(x+1),W^{(i)},X_i^*)\Big],\end{aligned} \tag{11}$$

where for any integer $d\ge 1$ and any $\mathbf{J}\in\mathbb{N}_*^d$, $\mathbf{J}^\dagger\in\mathbb{N}_*$ denotes the last coordinate of $\mathbf{J}$, and $\mathbf{J}^\circ$ denotes the element in $\mathbb{N}_*^{d-1}$ obtained from $\mathbf{J}$ by omitting the last coordinate.

Remark 1.3 In view of the similarity between the above theorem and [12, Thm1.2], which has also been shown by the two papers [2, 3] of Barbour, the following question arises naturally: can we generalize the result to any infinitely divisible distribution?

2 Several preliminary results

In this section, we are interested in some properties concerning the function h and the associated function f_h. Compared to the normal case, we no longer need differentiability conditions on h in Theorem 1.2 and shall concentrate on its increasing speed at infinity. This makes the study much simpler.

We begin by considering the modified Stein's equation on $\mathbb{N}_*$:

$$x\widetilde{f}(x) - \lambda\widetilde{f}(x+1) = h(x), \qquad x \in \mathbb{N}_*. \tag{12}$$

The above equation may have many solutions, one of which is given by

$$\widetilde{f_h}(x) := \frac{(x-1)!}{\lambda^x}\sum_{i=x}^{\infty}\frac{\lambda^i}{i!}h(i).$$

A general solution of (12) can be written as $\widetilde{f_h}(x) + C(x-1)!/\lambda^x$, where C is an arbitrary constant. However, when h is of polynomial increasing speed at infinity, $\widetilde{f_h}$ is the only solution of (12) which has polynomial increasing speed at infinity.

In order that the function $\widetilde{f_h}$ is well defined, we need some condition on h. Denote by $\mathcal{E}_\lambda$ the space of functions h on $\mathbb{N}_*$ such that, for any polynomial P, we have

$$\sum_{i\geq 1}\frac{\lambda^i}{i!}|h(i)P(i)| < +\infty.$$

Clearly $\mathcal{E}_\lambda$ is a linear space. We list below some properties of $\mathcal{E}_\lambda$.

Proposition 2.1 *The following assertions hold:*

1) for any $Q \in \mathbb{R}[x, x^{-1}]$ and any $h \in \mathcal{E}_\lambda$ where $\mathbb{R}[x, x^{-1}]$ denotes the set of Laurent polynomials on $\mathbb{R}$, we have $Qh \in \mathcal{E}_\lambda$;

2) for any $h \in \mathcal{E}_\lambda$, $\Delta h \in \mathcal{E}_\lambda$ and $\widetilde{f_h} \in \mathcal{E}_\lambda$;

Proof. 1) is obvious by definition.
2) Let h_1 be the function defined as $h_1(x) := h(x+1)$. If P is a polynomial, then

$$\sum_{i\geq 1}\frac{\lambda^i}{i!}\big|P(i)h_1(i)\big| = \sum_{j\geq 2}\frac{\lambda^{j-1}}{(j-1)!}\big|P(j-1)h(j)\big| = \lambda^{-1}\sum_{j\geq 2}\frac{\lambda^j}{j!}\big|jP(j-1)h(j)\big| < +\infty$$

since $h \in \mathcal{E}_\lambda$. Therefore, $\Delta h \in \mathcal{E}_\lambda$. We next prove the second assertion. For any arbitrary polynomial P, there exists another polynomial Q such that, for any integer $i \geq 1$, $Q(i) \geq \sum_{j=1}^{i} \frac{|P(j)|}{j}$. Therefore

$$\sum_{a\geq 1} \frac{\lambda^a}{a!}\Big|P(a)\frac{(a-1)!}{\lambda^a}\sum_{i=a}^{\infty}\frac{\lambda^i}{i!}h(i)\Big| \leq \sum_{a\geq 1}\frac{|P(a)|}{a}\sum_{i=a}^{\infty}\frac{\lambda^i}{i!}|h(i)|$$

$$= \sum_{i\geq 1}\Big(\sum_{a=1}^{i}\frac{|P(a)|}{a}\Big)\frac{\lambda^i}{i!}|h(i)| \leq \sum_{i\geq 1} Q(i)\frac{\lambda^i}{i!}|h(i)| < +\infty,$$

which implies that $\widetilde{f}_h \in \mathcal{E}_\lambda$. □

For any function $h \in \mathcal{E}_\lambda$, we define $\tau(h) : \mathbb{N} \to \mathbb{R}$ such that

$$\tau(h)(x) = h(x+1)/(x+1).$$

Note that for any integer $k \geq 1$, one has $\tau^k(h)(x) = x!h(x+k)/(x+k)!$. The proof of Proposition 2.1 shows that τ is actually an endomorphism of $\mathcal{E}_\lambda$.

Lemma 2.2 *Let $h \in \mathcal{E}_\lambda$. Then*

$$\widetilde{f}_{\tau(h)}(x) = \widetilde{f}_h(x+1)/x. \tag{13}$$

Proof. Let $u(x) = \widetilde{f}_h(x+1)/x$. Dividing both sides of (12) by x and then replacing x by $x+1$, we obtain $\widetilde{f}_h(x+1) - \lambda\widetilde{f}_h(x+2)/(x+1) = \tau(h)(x)$, or equivalently,

$$xu(x) - \lambda u(x+1) = \tau(h)(x). \tag{14}$$

Since $\widetilde{f}_{\tau(h)}$ is the only solution of (14) in $\mathcal{E}_\lambda$, the lemma is proved. □

Corollary 2.3 *Let $p \in \mathbb{R}$. If $h(x) = O(x^p)$, then $\widetilde{f}_h(x) = O(x^{p-1})$.*

Proof. First of all,

$$0 \leq \frac{x!}{\lambda^x}\sum_{i\geq x}\frac{\lambda^i}{i!} = \sum_{i\geq x}\frac{\lambda^{i-x}}{i!/x!} \leq \sum_{i\geq x}\frac{\lambda^{i-x}}{(i-x)!} = e^{-\lambda}.$$

Therefore, when $p \leq 0$, one has

$$x\widetilde{f}_h(x) = \frac{x!}{\lambda^x}\sum_{i=x}^{\infty}\frac{\lambda^i}{i!}h(i) = O(x^p)$$

since $i^p \leq x^p$ if $x \leq i$. Hence $\widetilde{f}_h(x) = O(x^{p-1})$. The general case follows by induction on p by using (13). □

We now introduce the function space: for any $p \in \mathbb{R}$, denote by $\mathcal{H}_p$ the space of all functions $h : \mathbb{N} \to \mathbb{R}$ such that $h(x) = O(x^p)$ when $x \to \infty$. In the following are some simple properties of $\mathcal{H}_p$, their proofs are direct.

Proposition 2.4 *1) For any $p \geq 0$ and any $h \in \mathcal{H}_p$, the restriction of h on $\mathbb{N}_*$ lies in $\bigcap_{\lambda>0} \mathcal{E}_\lambda$.*

2) If $h \in \mathcal{H}_p$, then also are $h(x+1)$ and Δh.

3) If $h \in \mathcal{H}_p$ and $g \in \mathcal{H}_q$, then $gh \in \mathcal{H}_{p+q}$.

The following proposition is essential for applying the recursive estimation procedure.

Proposition 2.5 *Let $p \geq 0$. If $h \in \mathcal{H}_p$, then $f_h \in \mathcal{H}_p$.*

Proof. Note that f_h coincides with $\widetilde{f_{\overline{h}}}$ on $\mathbb{N}_*$ where $\overline{h} = h - \mathcal{P}_\lambda(h)$. Since $h \in \mathcal{H}_p$, also is $\overline{h}$. Then Corollary 2.3 implies $f_h(x) = O(x^{p-1}) = O(x^p)$. □

3 Proof of the main result

In this section, we give the proof of Proposition 1.1 and of Theorem 1.2, which are essentially the same with the ones of [12, Prop1.1, Thm1.2] in a discrete setting.

Proof of Proposition 1.1 We replace $\mathbb{E}[\Delta^{|\mathbf{J}|} f(X+Y)]$ on the right side of (8) by

$$\sum_{k=0}^{N-|\mathbf{J}|} m_Y^{(k)} \mathbb{E}[\Delta^{|\mathbf{J}|+k} f(X)] + \delta_{N-|\mathbf{J}|}(\Delta^{|\mathbf{J}|} f, X, Y)$$

and observe that the sum of terms containing δ vanishes with $\varepsilon_N(f, X, Y)$. Hence the right side of (8) equals

$$\sum_{d\geq 0} (-1)^d \sum_{\mathbf{J}\in\mathbb{N}_*^d,\, |\mathbf{J}|\leq N} m_Y^{(\mathbf{J})} \sum_{k=0}^{N-|\mathbf{J}|} m_Y^{(k)} \mathbb{E}[\Delta^{|\mathbf{J}|+k} f(X)]$$

If we split the terms for $k = 0$ and for $1 \leq k \leq N - |\mathbf{J}|$ respectively, the above formula can be written as

$$(15) \quad \sum_{d\geq 0} (-1)^d \sum_{\mathbf{J}\in\mathbb{N}_*^d,\, |\mathbf{J}|\leq N} m_Y^{(\mathbf{J})} \mathbb{E}[\Delta^{|\mathbf{J}|} f(X)] + \sum_{d\geq 0} (-1)^d \sum_{\mathbf{J}\in\mathbb{N}_*^d,\, |\mathbf{J}|\leq N} m_Y^{(\mathbf{J})} \sum_{k=1}^{N-|\mathbf{J}|} m_Y^{(k)} \mathbb{E}[\Delta^{|\mathbf{J}|+k} f(X)].$$

We make the index changes $\mathbf{J}' = (\mathbf{J}, k)$ and $u = d + 1$ in the second part of (15) and find that it is nothing but

$$\sum_{u\geq 1} (-1)^{u-1} \sum_{\mathbf{J}'\in\mathbb{N}_*^u,\, |\mathbf{J}'|\leq N} m_Y^{(\mathbf{J}')} \mathbb{E}[\Delta^{|\mathbf{J}'|} f(X)].$$

By taking the sum, it only remains the term of index $d = 0$ in the first part of (15), which is equal to $\mathbb{E}[f(X)]$. So the lemma is proved. $\square$

Proof of Theorem 1.2 We prove the theorem by induction on N. The case where $N = 0$ is trivial. Assume that the assertion holds for $0, \cdots, N-1$.
Since $h \in \mathcal{H}_p$, by Lemma 2.5 and Proposition 2.4 2), for any $k \in \{1, \cdots, N\}$, $\Delta^k f_h(x+1) \in \mathcal{H}_{p-1} \subset \mathcal{H}_p$. Therefore $C_{N-k}(\Delta^k f_h(x+1))$ and $e_{N-k}(\Delta^k f_h(x+1))$ are well defined and

$$\mathbb{E}[\Delta^k f_h(W+1)] = C_{N-k}(\Delta^k f_h(x+1)) + e_{N-k}(\Delta^k f_h(x+1)).$$

We now prove the equality $\mathbb{E}[h(W)] = C_N(h) + e_N(h)$. Recall that for any $i \in \{1, \cdots, n\}$, X_i^* follows the Poisson X_i-zero biased distribution and is independent of $W^{(i)} = W - X_i$, I is an independent random index such that $\mathbb{P}(I = i) = \lambda_i/\lambda_W$, and $W^* = W^{(I)} + X_I^*$. So $\mathbb{E}[h(W)] - C_0(h)$ is equal to

$$\lambda_W \mathbb{E}[f_h(W^*+1) - f_h(W+1)] = \sum_{i=1}^n \lambda_i \Big(\mathbb{E}[f_h(W^{(i)} + X_i^* + 1)] - \mathbb{E}[f_h(W+1)] \Big),$$

where, by using (6),

$$\mathbb{E}[f_h(W^{(i)} + X_i^* + 1)] = \sum_{k=0}^N m_{X_i^*}^{(k)} \mathbb{E}[\Delta^k f_h(W^{(i)} + 1)] + \delta_N(f_h(x+1), W^{(i)}, X_i^*).$$

By replacing $\mathbb{E}[\Delta^k f_h(W^{(i)}+1)]$ in the above formula by its $(N-k)^{\text{th}}$ order reverse Taylor expansion, we obtain that $\mathbb{E}[f_h(W^{(i)} + X_i^* + 1)]$ equals

$$\sum_{k=0}^N m_{X_i^*}^{(k)} \Bigg[\sum_{d \geq 0} (-1)^d \sum_{\substack{\mathbf{J} \in \mathbb{N}_*^d \\ |\mathbf{J}| \leq N-k}} m_{X_i}^{(\mathbf{J})} \mathbb{E}[\Delta^{|\mathbf{J}|+k} f_h(W+1)] + \varepsilon_{N-k}(\Delta^k f_h(x+1), W^{(i)}, X_i) \Bigg] + \delta_N(f_h(x+1), W^{(i)}, X_i^*).$$

Note that the term of indices $k = d = 0$ in the sum is $\mathbb{E}[f_h(W+1)]$. Therefore, $\mathbb{E}[f_h(W^{(i)} + X_i^* + 1)] - \mathbb{E}[f_h(W+1)]$ is the sum of the following three terms

$$\sum_{k=1}^N m_{X_i^*}^{(k)} \sum_{d \geq 0} (-1)^d \sum_{\mathbf{J} \in \mathbb{N}_*^d,\, |\mathbf{J}| \leq N-k} m_{X_i}^{(\mathbf{J})} \mathbb{E}[\Delta^{|\mathbf{J}|+k} f_h(W+1)], \tag{16}$$

$$\sum_{d \geq 1} (-1)^d \sum_{\mathbf{J} \in \mathbb{N}_*^d,\, |\mathbf{J}| \leq N} m_{X_i}^{(\mathbf{J})} \mathbb{E}[\Delta^{|\mathbf{J}|} f_h(W+1)], \tag{17}$$

$$\sum_{k=0}^N m_{X_i^*}^{(k)} \varepsilon_{N-k}(\Delta^k f_h(x+1), W^{(i)}, X_i) + \delta_N(f_h(x+1), W^{(i)}, X_i^*). \tag{18}$$

By interchanging summations and then making the index changes $\mathbf{K} = (\mathbf{J}, k)$ and $u = d+1$, we obtain

$$(16) = \sum_{u \geq 1} (-1)^{u-1} \sum_{\mathbf{K} \in \mathbb{N}_*^u,\, |\mathbf{K}| \leq N} m_{X_i}^{(\mathbf{K}^\circ)} m_{X_i^*}^{(\mathbf{K}^\dagger)} \mathbb{E}[\Delta^{|\mathbf{K}|} f_h(W+1)].$$

As the equality $m_{X_i}^{(\mathbf{J})} = m_{X_i}^{(\mathbf{J}^\circ)} m_{X_i}^{\mathbf{J}^\dagger}$ holds for any $\mathbf{J}$, (16)+(17) simplifies as

$$\sum_{d\geq 1}(-1)^{d-1}\sum_{\mathbf{J}\in\mathbb{N}_*^d,\,|\mathbf{J}|\leq N} m_{X_i}^{(\mathbf{J}^\circ)}\Big(m_{X_i^*}^{(\mathbf{J}^\dagger)} - m_{X_i}^{(\mathbf{J}^\dagger)}\Big)\mathbb{E}[\Delta^{|\mathbf{J}|}f_h(W+1)].$$

By the hypothesis of induction, we have

$$\mathbb{E}[\Delta^{|\mathbf{J}|}f_h(W+1)] = C_{N-|\mathbf{J}|}(\Delta^{|\mathbf{J}|}f_h(x+1)) + e_{N-|\mathbf{J}|}(\Delta^{|\mathbf{J}|}f_h(x+1)),$$

so the equality $\mathbb{E}[h(W)] = C_N(h) + e_N(h)$ holds with $C_N(h)$ and $e_N(h)$ being defined in (10) and (11).

□

4 Error estimations

In this section, we concentrate on the remainder $e_N(h)$ in the asymptotic expansion. The following quantity will be useful. Let $p \geq 0$. For $h \in \mathcal{H}_p$ and $N \in \mathbb{N}$, we define

$$\|h\|_{N,p} := \sup_{x\in\mathbb{N}_*} \frac{\left|\Delta^{N+1}h(x)\right|}{x^p}, \tag{19}$$

which is finite by Proposition 2.4 2).

Lemma 4.1 *Let $N \in \mathbb{N}$, $k \in \{0,\cdots,N\}$ and $p \geq 0$. Let X be an $\mathbb{N}$-r.v. with p^{th} order moment, Y be an $\mathbb{N}$-r.v. independent of X and having $(N-k+1+p)^{th}$ order moment. Then, for any $f \in \mathcal{H}_p$, the following inequalities hold:*

$$|\delta_{N-k}(\Delta^k f(x+1), X, Y)| \leq \max(2^{p-1},1)\|f\|_{N,p}\big(\mathbb{E}[X^p]m_Y^{(N-k+1)} + m_Y^{(N-k+1),p}\big), \tag{20}$$

where

$$m_Y^{(N-k+1),p} := \mathbb{E}\Big[\binom{Y}{N-k+1}Y^p\Big].$$

The discrete reverse Taylor remainder satisfies

$$\begin{aligned}&\left|\varepsilon_{N-k}(\Delta^k f(x+1),X,Y)\right| \\ &\leq \max(2^{p-1},1)\|f\|_{N,p}\sum_{d\geq 0}\sum_{\substack{\mathbf{J}\in\mathbb{N}_*^d\\|\mathbf{J}|\leq N-k}} m_Y^{(\mathbf{J})}\big(\mathbb{E}[X^p]m_Y^{(N-k-|\mathbf{J}|+1)} + m_Y^{(N-k-|\mathbf{J}|+1),p}\big).\end{aligned} \tag{21}$$

Proof. By definition (7) and (19),

$$\begin{aligned}&\left|\delta_{N-k}(\Delta^k f(x+1),X,Y)\right| \leq \mathbb{E}\Big[\sum_{0\leq j_1<\cdots<j_{N-k+1}<Y}\left|\Delta^{N+1}f(X+1+j_1)\right|\Big] \\ &\leq \|f\|_{N,p}\,\mathbb{E}\Big[\sum_{0\leq j_1<\cdots<j_{N-k+1}<Y}(X+j_1+1)^p\Big] \leq \|f\|_{N,p}\,\mathbb{E}\Big[\binom{Y}{N-k+1}(X+Y)^p\Big] \\ &\leq \max(2^{p-1},1)\|f\|_{N,p}\big(\mathbb{E}[X^p]m_Y^{(N-k+1)} + m_Y^{(N-k+1),p}\big),\end{aligned}$$

where we have used in the last inequality the estimations $(X+Y)^p \leq 2^{p-1}(X^p+Y^p)$ if $p>1$ and $(X+Y)^p \leq X^p+Y^p$ if $p \leq 1$. Thus (20) is proved. The inequality (21) follows from (9) and (20). □

Proposition 4.2 *Let $N \in \mathbb{N}$, $p \geq 0$ and $h \in \mathcal{H}_p$. Let $X_i\,(i=1,\cdots,n)$ be a family of independent $\mathbb{N}$-r.v.s with mean $\lambda_i > 0$ and up to $(N+p+1)^{th}$ order moments; $W = X_1+\cdots+X_n$. Let X_i^* be an $\mathbb{N}$-r.v. having Poisson X_i-zero biased distribution and independent of $W^{(i)} := W - X_i$. Then the following estimations hold.*

1) When $N=0$,

(22)

$$|e_0(h)| \leq \max(2^{p-1},1)\|f_h\|_{0,p}\sum_{i=1}^n \Big(\mathbb{E}[(W^{(i)})^p]\big(\mathbb{E}[X_i^2]+\lambda_i^2-\lambda_i\big)+\lambda_i\big(\mathbb{E}[(X_i^*)^{p+1}]+\mathbb{E}[(X_i)^{p+1}]\big)\Big).$$

2) When $N \geq 1$, one has the recursive estimation:

(23)

$$|e_N(h)| \leq \sum_{i=1}^n \lambda_i\Big[\sum_{d\geq 1}\sum_{\substack{\mathbf{J}\in\mathbb{N}_*^d\\ |\mathbf{J}|\leq N}} m_{X_i}^{(\mathbf{J}^\circ)}\big(m_{X_i^*}^{(\mathbf{J}^\dagger)}+m_{X_i}^{(\mathbf{J}^\dagger)}\big)|e_{N-|\mathbf{J}|}(\Delta^{|\mathbf{J}|}f_h(x+1))|$$

$$+\max(2^{p-1},1)\|f_h\|_{N,p}\sum_{k=0}^N m_{X_i^*}^{(k)}\sum_{d\geq 0}\sum_{\substack{\mathbf{J}\in\mathbb{N}_*^d\\ |\mathbf{J}|\leq N-k}} m_{X_i}^{(\mathbf{J})}\big(\mathbb{E}[(W^{(i)})^p]m_{X_i}^{(N-k-|\mathbf{J}|+1)}+m_{X_i}^{(N-k-|\mathbf{J}|+1),p}\big)$$

$$+\max(2^{p-1},1)\|f_h\|_{N,p}\big(\mathbb{E}[(W^{(i)})^p]m_{X_i^*}^{(N+1)}+m_{X_i^*}^{(N+1),p}\big)\Big],$$

Proof. We begin by the case when $N=0$. By (4),

$$\begin{aligned} e_0(h) &= \sum_{i=1}^n \lambda_i\big(\mathbb{E}[f_h(W^{(i)}+X_i^*+1)]-\mathbb{E}[f_h(W^{(i)}+X_i+1)]\big)\\ &= \sum_{i=1}^n \lambda_i\big(\delta_0(f_h(x+1),W^{(i)},X_i^*)+\varepsilon_0(f_h(x+1),W^{(i)},X_i)\big)\\ &\leq \max(2^{p-1},1)\|f_h\|_{0,p}\sum_{i=1}^n \lambda_i\Big\{\mathbb{E}[(W^{(i)})^p]\big(m_{X_i^*}^{(1)}+m_{X_i}^{(1)}\big)+\big(m_{X_i^*}^{(1),p}+m_{X_i}^{(1),p}\big)\Big\} \end{aligned}$$

where the last inequality is by estimations (20) and (21), so (22) follows. Combining in addition the recursive formula (11), we obtain the inequality (23). □

References

[1] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *The Annals of Probability*, 17(1):9–25, 1989.

[2] A. D. Barbour. Asymptotic expansions based on smooth functions in the central limit theorem. *Probability Theory and Related Fields*, 72:289–303, 1986.

[3] A. D. Barbour. Asymptotic expansions in the Poisson limit theorem. *Annals of Probability*, 15(2):748–766, 1987.

[4] A. D. Barbour and V. Čekanavičius. Total variation asymptotics for sums of independent integer random variables. *The Annals of Probability*, 30(2):509–545, 2002.

[5] A. D. Barbour, L. H. Y. Chen, and K. P. Choi. Poisson approximation for unbounded functions. I. Independent summands. *Statistica Sinica*, 5(2):749–766, 1995.

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

[7] I. S. Borisov and P. S. Ruzankin. Poisson approximation for expectations of unbounded functions of independent random variables. *Annals of Probability*, 30(4):1657–1680, 2002.

[8] L. H. Y. Chen. Poisson approximation for dependent trials. *Annals of Probability*, 3:534–545, 1975.

[9] N. El Karoui and Y. Jiao. Stein's method and zero bias transformation for CDOs tranches pricing. *Finance and Stochastics*, 13(2):151–180, 2009.

[10] T. Erhardsson. Stein's method for Poisson and compound Poisson approximation. In *An Introduction to Stein's Method*, volume 4 of *Lecture Notes Series, IMS, National University of Singapore*, pages 61–113. Singapore University Press and World Scientific Publishing Co. Pte. Ltd., 2005.

[11] L. Goldstein and G. Reinert. Stein's method and the zero bias transformation with application to simple random sampling. *Annals of Applied Probability*, 7:935–952, 1997.

[12] Y. Jiao. Zero bias transformation and asymptotic expansions. Preprint, PMA Université Paris 6 - Université Paris 7, 2009.