The Annals of Probability
1987, Vol. 15, No. 2, 748–766

ASYMPTOTIC EXPANSIONS IN THE POISSON LIMIT THEOREM

BY A. D. BARBOUR

Universität Zürich

Asymptotic expansions for the distributions of sums of independent nonnegative integer random variables in the neighbourhood of the Poisson distribution are derived, together with explicit estimates of the truncation error. Expansions are also derived for the expectations of at most polynomially growing functions of such sums. Applications to the Poisson binomial and Poisson negative binomial approximations are considered. The method used is an adaptation of the Stein–Chen approach.

1. Introduction. In Chen (1975), Stein's method of obtaining error bounds for normal approximations was introduced in the Poisson context, and was used to obtain rates of convergence in total variation to the Poisson distribution for sums of stationary sequences of 0–1 random variables. Chen also observed that Stein's method is in principle suited to developing asymptotic expansions as well as obtaining error estimates, and used it for independent 0–1 summands to derive the second term in such an expansion, together with an estimate of the remaining error: Kerstan (1964) had previously derived a similar result, with a sharper error estimate, by a quite different technique. The Stein–Chen method was refined in Barbour and Hall (1984) to yield the best known upper bounds for the discrepancy between the Poisson distribution and that of a sum of independent 0–1 random variables, as well as complementary lower bounds, and new estimates for the error remaining after the second term in the asymptotic expansion were also established. However, although Stein's method was potentially applicable also to higher-order expansions, it proved in practice too cumbersome to use.

In this paper, it is shown how, by means of a simple identity, the Stein–Chen method can be made to yield a full asymptotic expansion with a minimum of difficulty. The error estimates obtained perform well in comparison with other estimates that are available. For sums of independent 0–1 random variables, they agree with those of Barbour and Hall, when the series is truncated after one term or after two. For the more general case of sums of independent nonnegative integer valued random variables, the error estimates obtained after two terms in the expansion differ only slightly from those in Barbour and Hall. Comparison with Kerstan's estimates, for integral nonnegative summands, of the error after one term of the expansion is more complicated. By and large, the techniques of this paper give better approximations, except when the summands themselves are almost precisely Poisson distributed. In Section 3, expansions are also obtained for the expectations of at most polynomially growing functions of the

Received February 1985; revised April 1986.
AMS 1980 *subject classifications.* 60G50, 60F05.
Key words and phrases. Poisson approximation, asymptotic expansions, nonnegative summands.

sum, which can be used to derive nonuniform estimates of the rate of convergence. Analogous results in the context of normal approximation are given in Barbour (1986).

Deheuvels and Pfeifer (1986) have recently used an elegant operator technique to study the error after one and two terms of the expansion in the case of 0–1 summands, but the only error estimates they give which are comparable with those obtained here are those of Kerstan. They are mostly concerned with asymptotic error estimates, valid when the mean of the distribution to be approximated tends to infinity. This reduces the problem to a study of the contribution of the first and second terms in the Poisson–Charlier measure defined in (2.7) below, rather than of the whole discrepancy.

The principal tool in the argument is Lemma 1.1 below. Let X be any nonnegative integer valued random variable, and let $m_{[j]}$ and $\kappa_{[j]}$ denote its jth factorial moment and cumulant, respectively. That is, $m_{[j]} := E\{X(X-1)\cdots(X-j+1)\}$ is the jth derivative of $E(z^X)$ at $z=1$, and $\kappa_{[j]}$ the jth derivative of $\log E(z^X)$. Set

$$\nu_l := \max\Big\{m_{[l+1]};\ \max_{0\le s\le l-1}\big(m_{[l-s]}|\kappa_{[s+1]}|\big)\Big\}.$$

Note that ν_l is a maximum over quantities of degree $l+1$ as moments of X, and replaces the $\mathbb{E}|X|^{l+1}$ error estimate available when ordinary, instead of factorial, moments and cumulants are being used. Note also that

$$m_{[l+1]} = \sum_{s=0}^{l}\binom{l}{s}m_{[l-s]}\kappa_{[s+1]}. \tag{1.1}$$

[Kendall and Stuart (1963), Section 3.17 and Example 3.9].

For any function $g\colon \mathbb{Z}^+ \to \mathbb{R}$, define

$$v_k(g;r) := \begin{cases} \max\limits_{0\le j\le r-k}|\Delta^k g(j)|, & r\ge k,\\ 0, & 0\le r<k,\end{cases}$$

and let g_j be the function defined by $g_j(x) := g(x+j)$.

LEMMA 1.1. *Suppose that, for some* $l\in\mathbb{N}$,

$$\mathbb{E}\{X^{l+1}v_l(g;X)\} < \infty.$$

Then

$$\mathbb{E}\{Xg(X)\} = \sum_{s=0}^{l-1}\frac{\kappa_{[s+1]}}{s!}\mathbb{E}\{\Delta^s g(X+1)\} + \eta_{l-1}(g), \tag{1.2}$$

where

$$|\eta_{l-1}(g)| \le \mathbb{E}\Big\{\max_{1\le j\le X}|\Delta^l g(j)|\psi_{l,1}(X)\Big\}, \tag{1.3}$$

and

$$\psi_{l,p}(X) := X^{p-1}\left\{X\binom{X-l}{l} + \sum_{s=0}^{l-1}\frac{|\kappa_{[s+1]}|}{s!}\binom{X}{l-s}\right\}.$$

REMARK. The conclusion of the lemma may be equivalently rewritten as

$$\mathbb{E}\{Xg(X) - (\mathbb{E}X)g(X+1)\} = \sum_{s=1}^{l-1} \frac{\kappa_{[s+1]}}{s!}\mathbb{E}\{\Delta^s g(X+1)\} + \eta_{l-1}(g).$$

PROOF. From Newton's interpolation formula, for $0 \le s \le l-1$,

$$\left|\Delta^s f(x) - \sum_{r=0}^{l-1-s} \binom{x}{r}\Delta^{r+s} f(0)\right| \le \binom{x}{l-s} \max_{0 \le j \le x-l+s} |\Delta^l f(j)|,$$

for any function f and $x \in \mathbb{Z}^+$. Hence

$$\left|\mathbb{E}(Xg(X)) - \sum_{r=0}^{l-1} \mathbb{E}\left\{X\binom{X-1}{r}\right\}\Delta^r g(1)\right| \le \mathbb{E}\left\{X\binom{X-1}{l}v_l(g_1; X)\right\} \tag{1.4}$$

and

$$\left|\mathbb{E}(\Delta^s g(X+1)) - \sum_{r=0}^{l-1-s} \mathbb{E}\left\{\binom{X}{r}\right\}\Delta^{r+s} g(1)\right| \le \mathbb{E}\left\{\binom{X}{l-s}v_{l-s}(\Delta^s g_1; X)\right\}, \tag{1.5}$$

$$0 \le s \le l-1.$$

But

$$\sum_{r=0}^{l-1} \mathbb{E}\left\{X\binom{X-1}{r}\right\}\Delta^r g(1) = \sum_{s=0}^{l-1} \frac{\kappa_{[s+1]}}{s!} \sum_{r=0}^{l-1-s} \mathbb{E}\left\{\binom{X}{r}\right\}\Delta^{r+s} g(1),$$

because of (1.1), and the lemma thus follows from (1.4) and (1.5). □

The next two corollaries follow immediately from estimate (1.3). Let $\phi_l := \mathbb{E}\psi_{l,1}(X)$, and note that $\phi_l \le 2^l v_l / l!$

COROLLARY 1.2. *If, for some* $l \in \mathbb{N}$, $\|\Delta^l g\| := \sup_{j \ge 0}|\Delta^l g(j)| < \infty$ *and* $\mathbb{E}X^{l+1} < \infty$, *then, for all* $j \ge 0$,

$$\left|\mathbb{E}\{Xg_j(X)\} - \sum_{s=0}^{l-1} \frac{\kappa_{[s+1]}}{s!}\mathbb{E}\{\Delta^s g_j(X+1)\}\right| \le \phi_l \|\Delta^l g\| \le 2^l v_l \|\Delta^l g\| / l!.$$

COROLLARY 1.3. *If* X *is a* 0–1 *random variable and* $\|\Delta^l g\| < \infty$, *for some* $l \in \mathbb{N}$, *then, for all* $j \ge 0$,

$$\left|\mathbb{E}\{Xg_j(X)\} - \sum_{s=0}^{l-1} (-1)^s p^{s+1}\mathbb{E}\{\Delta^s g_j(X+1)\}\right| \le p^{l+1}\|\Delta^l g\|,$$

where $p := P[X = 1]$.

COROLLARY 1.4. *Suppose* X *has a factorial moment generating function with nonzero radius of convergence, and let* g *satisfy*

$$\sup_{j \ge 0} |\Delta^k g(j)| \le Cr^k, \qquad k \ge m, \tag{1.6}$$

for some $C, m > 0$ and $r < R$, where R denotes the radius of convergence of the factorial cumulant generation function of X. Then the identity

$$\mathbb{E}\{Xg(X)\} = \sum_{s\geq 0} \frac{\kappa_{[s+1]}}{s!}\mathbb{E}\{\Delta^s g(X+1)\} \tag{1.7}$$

holds.

PROOF. The radius of convergence of the factorial moment generating function of X is no less than R. Hence, for $l \geq m$,

$$\begin{aligned}|\eta_{l-1}(g)| &\leq C' r^l \left\{ \frac{(l+1)}{(r')^{l+1}} + \sum_{s=0}^{l-1} \frac{(s+1)}{(r')^{s+1}} \frac{1}{(r')^{l-s}} \right\} \\ &\leq C' l^2 \left(\frac{r}{r'}\right)^l\end{aligned}$$

for any $r < r' < R$ and for suitable constants C'. Hence

$$\lim_{l\to\infty} |\eta_{l-1}(g)| = 0. \qquad \square$$

REMARK. If $\Delta^m g$ is bounded, it follows automatically that (1.6) holds with $r = 2$.

2. Expansions for distributions. In this section, it is assumed throughout that $(X_i)_{i=1}^N$ are independent nonnegative integer valued random variables, the distribution of whose sum W is to be approximated. The notation of Section 1 is carried over in the natural way with $m_{[j]}^{(i)}$, $\kappa_{[j]}^{(i)}$, $\phi_j^{(i)}$ and $\nu_j^{(i)}$ referring to the X_i-distribution: define also $\kappa_{[j]} := \sum_{i=1}^N \kappa_{[j]}^{(i)}$, the jth factorial cumulant of W, $\phi_j := \sum_{i=1}^N \phi_j^{(i)}$, $\nu_j := \sum_{i=1}^N \nu_j^{(i)}$, $\lambda := \mathbb{E}W$ and $W_i := W - X_i$. The symbol Q is used throughout to denote a random variable with the Poisson distribution P_λ with mean λ.

Then the following lemma is a consequence of Corollaries 1.2 and 1.3:

LEMMA 2.1. *Suppose that $g\colon \mathbb{Z}^+ \to \mathbb{R}$ satisfies $\|\Delta^l g\| < \infty$ for some $l \in \mathbb{N}$. Then, if $\mathbb{E}X_i^{l+1} < \infty$, $1 \leq i \leq N$,*

$$\left|\mathbb{E}\{Wg(W) - \lambda g(W+1)\} - \sum_{s=1}^{l-1} \frac{1}{s!}\kappa_{[s+1]}\mathbb{E}\{\Delta^s g(W+1)\}\right| \leq \phi_l \|\Delta^l g\|. \tag{2.1}$$

PROOF. From Corollary 1.2,

$$\left|\mathbb{E}\{X_i g(W) - (\mathbb{E}X_i)g(W+1)\} - \sum_{s=1}^{l-1} \frac{\kappa_{[s+1]}^{(i)}}{s!}\mathbb{E}\{\Delta^s g(W+1)\}\right| \leq \phi_l^{(i)} \mathbb{E}\|\Delta^l g_{W_i}\|, \qquad 1 \leq i \leq \mathbb{N}, \tag{2.2}$$

and (2.1) follows by adding over i. $\square$

The first step in Chen's (1975) approximation technique is to observe that, for any function h: $\mathbb{Z}^+ \to \mathbb{R}$ such that $\bar{h} := \mathbb{E}h(Q)$ exists, the function $\theta_\lambda h$ defined by $\theta_\lambda h(-1) = 0$ and

$$\begin{aligned}\theta_\lambda h(m) &= m!\lambda^{-m-1}\sum_{j=0}^{m}\{\lambda^j(h(j)-\bar{h})/j!\} \\ &= -m!\lambda^{-m-1}\sum_{j=m+1}^{\infty}\{\lambda^j(h(j)-\bar{h})/j!\}, \qquad m \geq 0,\end{aligned} \tag{2.3}$$

satisfies

$$\lambda\theta_\lambda h(m) - m\theta_\lambda h(m-1) = h(m) - \bar{h}, \qquad m \geq 0. \tag{2.4}$$

Hence the estimate (2.1) of Lemma 2.1 directly yields an asymptotic expansion

$$\mathbb{E}h(W) - \mathbb{E}h(Q) = -\sum_{s=1}^{l-1}\frac{1}{s!}\kappa_{[s+1]}\mathbb{E}\{\Delta^s\theta_\lambda h(W)\} + \eta, \tag{2.5}$$

where, for functions h such that $\|\Delta^l\theta_\lambda h\| < \infty$,

$$|\eta| \leq \phi_l\|\Delta^l\theta_\lambda h\|, \tag{2.6}$$

from (2.1). By using the left-hand side of (2.5) successively to obtain expansions of the right-hand side of (2.5), an asymptotic expansion for $\mathbb{E}h(W)$ can be deduced.

To state the theorem, some extra notation is needed. Let the Charlier polynomials $C_n(\lambda; x)$ be defined by

$$C_n(\lambda; x) := \sum_{r=0}^{n}\binom{n}{r}(-1)^{n-r}\lambda^{-r}x_{[r]},$$

where $x_{[r]} := x(x-1)\cdots(x-r+1)$ denotes the rth factorial power of x.

Then define signed measures Q_l, $l \geq 1$, on $\mathbb{Z}^+$ by

$$Q_l(m) := \left\{\frac{e^{-\lambda}\lambda^m}{m!}\right\}\left\{1 + \sum_{s=1}^{l-1}\sum_{[s]}\prod_{j=1}^{s}\left\{\frac{1}{r_j!}\left(\frac{\kappa_{[j+1]}}{(j+1)!}\right)^{r_j}\right\}C_{R+s}(\lambda; m)\right\}, \tag{2.7}$$

where $\Sigma_{[s]}$ denotes the sum over all $(r_1, \ldots, r_s) \in (\mathbb{Z}^+)^s$ such that $\Sigma_{j=1}^s jr_j = s$ and $R := \Sigma_{j=1}^s r_j$.

Further, let

$$e_\lambda := \lambda^{-1}(1 - e^{-\lambda}) \leq \min(\lambda^{-1}, 1),$$

and set

$$\mu_l := \max_{(s)}\left\{e_\lambda^{k+1}\phi_{s_{k+1}}\prod_{j=1}^{k}\left|\frac{\kappa_{[s_j+1]}}{s_j!}\right|\right\},$$

$$\mu_l' := \max_{(s)}\left\{\frac{e_\lambda^{k+1}\nu_{s_{k+1}}}{(s_{k+1}-1)!}\prod_{j=1}^{k}\left|\frac{\kappa_{[s_j+1]}}{s_j!}\right|\right\},$$

where $\max_{(s)}$ denotes the maximum over $\{k \geq 0;\ s_j \geq 1,\ 1 \leq j \leq k+1;\ \Sigma_{j=1}^{k+1}s_j = l\}$.

THEOREM 1. *Suppose that* h: $\mathbb{Z}^+ \to \mathbb{R}$ *satisfies* $\|h\| < \infty$ *and that* $\mathbb{E}X_i^{l+1} < \infty$, $1 \le i \le N$; *then*

$$\mathbb{E}h(W) = \int h\,dQ_l + \eta_l, \tag{2.8}$$

where

$$|\eta_l| \le \begin{cases} 2^{2l-1}\mu_l\|h\|, \\ 2^{2l}\mu_l'\left(\dfrac{1}{l} + \dfrac{1}{2}\displaystyle\sum_{s=1}^{l-1}\dfrac{1}{s}\right)\|h\|. \end{cases} \tag{2.9}$$

REMARKS. 1. Inequality (2.12) below is sharper, but less convenient, than (2.9).

2. The use of truncation can extend the range application of Theorem 1. Thus, defining, for some $m \ge 1$,

$$Y_i = \begin{cases} X_i, & X_i \le m, \\ 0, & X_i > m, \end{cases}$$

and letting $Z = \Sigma_{i=1}^N Y_i$, it follows, as observed by Kerstan (1964) and Serfling (1975), that, if $\|h\| < \infty$, $|\mathbb{E}h(W) - \mathbb{E}h(Z)| \le 2\|h\|\Sigma_{i=1}^N P[X_i > m]$. Theorem 1 is then automatically available for the approximation of $\mathbb{E}h(Z)$.

In order to prove Theorem 1, we need the following results:

LEMMA 2.2. *Let* h: $\mathbb{Z}^+ \to \mathbb{R}$ *satisfy* $\|h\| < \infty$. *Then* $\|\Delta\theta_\lambda h\| \le 2e_\lambda\|h\|$.

PROOF. Define $h^+(j) := h(j) - \inf_k h(k) \ge 0$. Clearly, $\theta_\lambda h^+ = \theta_\lambda h$. But also,

$$\theta_\lambda h^+ = \sum_{j\ge 0} h^+(j)\theta_\lambda\delta_j, \tag{2.10}$$

where δ_j: $\mathbb{Z}^+ \to \mathbb{R}$ is defined by

$$\delta_j(m) = \begin{cases} 1, & j = m, \\ 0, & j \ne m. \end{cases}$$

From the argument used to prove Lemma 4(ii) of Barbour and Eagleson (1983), it now follows that

$$\|\Delta\theta_\lambda h\| \le e_\lambda\left\{\sup_{j\ge 0} h^+(j)\right\},$$

giving the required result. □

COROLLARY 2.3. *Let* $\|h\| < \infty$. *Then, for any choice of* $r, s_1, s_2, \ldots, s_r \ge 1$ *such that* $\Sigma_{j=1}^r s_j = l \ge 1$,

$$\left\|\prod_{j=1}^r (\Delta^{s_j}\theta_\lambda)h\right\| \le 2^l e_\lambda^r\|h\|.$$

PROOF. It is immediate, for any function g: $\mathbb{Z}^+ \to \mathbb{R}$ such that $\|g\| < \infty$, that $\|\Delta^r g\| \le 2^r \|g\|$. Thus

$$\left\| \prod_{j=1}^{r} (\Delta^{s_j}\theta_\lambda)h \right\| \le 2^{s_r - 1} \left\| \Delta\theta_\lambda \prod_{j=1}^{r-1} (\Delta^{s_j}\theta_\lambda)h \right\|$$

$$\le 2^{s_r} e_\lambda \left\| \prod_{j=1}^{r-1} (\Delta^{s_j}\theta_\lambda)h \right\|,$$

using Lemma 2.2; repeating the process, the corollary follows. □

PROOF OF THEOREM 1. Iteration of the relation (2.5) leads directly to the formula

$$\mathbb{E}h(W) = \sum_{(s)} (-1)^k \left(\prod_{j=1}^{k} \frac{\kappa_{[s_j+1]}}{s_j!} \right) \mathbb{E}\left\{ \prod_{j=1}^{k} (\Delta^{s_j}\theta_\lambda)h(Q) \right\} + \eta_l, \tag{2.11}$$

where

$$|\eta_l| \le \sum_{(s)} \left| \prod_{j=1}^{k} \frac{\kappa_{[s_j+1]}}{s_j!} \right| \phi_{s_{k+1}} \left\| \prod_{j=1}^{k+1} (\Delta^{s_j}\theta_\lambda)h \right\|;$$

here, $\Sigma_{(s)}$ denotes the sum over $\{k \ge 0;\ s_j \ge 1,\ 1 \le j \le k+1;\ \Sigma_{j=1}^{k+1} s_j = l\}$.

It follows immediately from Corollary 2.3 that

$$|\eta_l| \le 2^l \sum_{(s)} \left| \prod_{j=1}^{k} \frac{\kappa_{[s_j+1]}}{s_j!} \right| e_\lambda^{k+1} \phi_{s_{k+1}} \|h\|, \tag{2.12}$$

whence

$$|\eta_l| \le 2^l \|h\| \mu_l \sum_{(s)} 1 = 2^{2l-1} \mu_l \|h\|,$$

and

$$|\eta_l| \le 2^l \|h\| \mu_l' \sum_{(s)} s_{k+1}^{-1} 2^{s_{k+1}}$$

$$= 2^{2l} \mu_l' \left\{ \frac{1}{l} + \frac{1}{2} \sum_{s=1}^{l-1} \frac{1}{s} \right\} \|h\|.$$

The proof therefore merely requires the identification of (2.8) and (2.11). However, it follows from the properties of the Charlier polynomials with respect to the Poisson distribution that, for any function f: $\mathbb{Z}^+ \to \mathbb{R}$ for which the expectations below exist,

$$\mathbb{E}\{C_n(\lambda; Q)(\Delta^m f)(Q)\} = \mathbb{E}\{C_{n+m}(\lambda; Q) f(Q)\}, \qquad n, m \ge 0, \tag{2.13}$$

and

$$\mathbb{E}\{C_n(\lambda; Q)(\theta_\lambda f)(Q)\} = -\frac{1}{n+1} \mathbb{E}\{C_{n+1}(\lambda; Q) f(Q)\}, \qquad n \ge 0. \tag{2.14}$$

Thus

$$\mathbb{E}\left\{\prod_{j=1}^{k}(\Delta^{s_j}\theta_\lambda)h(Q)\right\} = (-1)^k\prod_{j=1}^{k}\left(j+\sum_{r=1}^{j}s_r\right)^{-1}\mathbb{E}\{C_{k+S}(\lambda;Q)h(Q)\},$$

where $S = \sum_{j=1}^{k}s_j$, and the equivalence of (2.8) and (2.11) follows from a standard combinatorial argument. □

In order to derive an asymptotic expansion for $P[W \in A]$, one uses $h(m) = I[m \in A] - \frac{1}{2}$, for which $\|h\| = \frac{1}{2}$. When $l = 1$, the estimate (2.9) of the discrepancy between $P[W \in A]$ and $P_\lambda(A)$ is then very similar to (4.6) of Theorem 4 in Barbour and Hall (1984). Also in the case $l = 1$, Kerstan (1964) proves that (2.8) holds with

$$|\eta_1| \le \min\{5.4\lambda^{-1}, 2\}\left(\sum_{i=1}^{N}d_i\right)\|h\|, \tag{2.15}$$

provided that $\max_{1\le i\le N}\{m_{[1]}^{(i)}\} \le \frac{1}{4}$, where

$$d_i = 2\Big\{\big(1 - m_{[1]}^{(i)} + \tfrac{1}{2}m_{[2]}^{(i)} - P[X_i = 0]\big) + \big(1 - m_{[1]}^{(i)} + \tfrac{1}{2}\big(m_{[1]}^{(i)}\big)^2 - \exp\big(-m_{[1]}^{(i)}\big)\big) + \big|\exp\big(-m_{[1]}^{(i)}\big) - P[X_i = 0]\big|\Big\}$$

$$\le m_{[2]}^{(i)} + \big(m_{[1]}^{(i)}\big)^2 = \phi_1^{(i)}.$$

Whether estimate (2.15) is better or worse than (2.12) with $l = 1$ depends on the distributions of the summands, though in both the particular cases considered below, of 0–1 and of negative binomial summands, (2.12) gives the better estimates. However, for Poisson distributed summands, where the discrepancy being estimated is zero, $\sum_{i=1}^{N}d_i \le \frac{2}{3}\sum_{i=1}^{N}(m_{[1]}^{(i)})^3$ is an order of magnitude smaller than the estimate (2.12).

An error estimate of similar order of magnitude can, however, be derived from (2.12), by taking $l = 2$, in the form

$$|\eta_2| \le 4e_\lambda\sum_{i=1}^{N}\Big\{m_{[1]}^{(i)}\big(\tfrac{1}{2}m_{[2]}^{(i)} + |\kappa_{[2]}^{(i)}|\big) + \tau_i + e_\lambda|\kappa_{[2]}|\Big(m_{[2]}^{(i)} + \big(m_{[1]}^{(i)}\big)^2\Big)\Big\}\|h\|, \tag{2.16}$$

where $\tau_i = \mathbb{E}\{X_i(X_i - 1)\min(\frac{1}{2}(X_i - 2), 1)\}$. The estimate

$$\mathbb{E}\left[\min\left\{X\binom{X-1}{l}v_l(g_1;X), 2X\binom{X-1}{l-1}v_{l-1}(g_1;X)\right\}\right]$$

of the remainder in (1.4) has been used in the $k = 0$ term $4e_\lambda\phi_2$, in order to avoid introducing moments of order three. For Poisson summands with means a_i, in Kerstan's range of application, this yields

$$|\eta_2| = |\eta_1| \le 4.5e_\lambda\sum_{i=1}^{N}a_i^3\|h\|,$$

still not quite as good as his estimate, in which $4.5e_\lambda$ is replaced by $\min(\frac{4}{3}, 3.6\lambda^{-1})$.

However, a very small perturbation away from the Poisson distribution for the summands, increasing $P[X_i = 0]$ and $P[X_i = 2]$ by $\frac{1}{3}a_i^3$ and reducing $P[X_i = 1]$ by $\frac{2}{3}a_i^3$, doubles Kerstan's estimate (2.15), while (2.16) increases by less than $10e_\lambda \sum_{i=1}^N a_i^4$; the estimate of error in (2.16) is then smaller in Kerstan's range of application than that of (2.15), for all $\lambda \leq 2.7$. Of course, (2.16) is then estimating the error in approximating $\mathbb{E}h(W)$ by $\int h\,dQ_2$ rather than by $\int h\,dQ_1$.

In the case of 0–1 summands, we have the following corollary:

COROLLARY 2.4. *If* $p_i = P[X_i = 1] = 1 - P[X_i = 0]$, $1 \leq i \leq N$, *then* (2.8) *holds with*

$$|\eta_l| \leq 2^{2l-1}e_\lambda\left(\sum_{i=1}^N p_i^{l+1}\right)\|h\|.$$

PROOF. Use estimate (2.12) and the evaluations

$$\kappa_{[m+1]} = (-1)^m m! \sum_{i=1}^N p_i^{m+1}, \qquad \phi_m = \sum_{i=1}^N p_i^{m+1},$$

to obtain

$$|\eta_l| \leq 2^l \sum_{(s)} \left\{\left[\prod_{j=1}^k \left(\sum_{i=1}^N p_i^{s_j+1}\right)\right]\sum_{i=1}^N p_i^{s_{k+1}+1}\right\} e_\lambda^{k+1}\|h\|. \tag{2.17}$$

Since the term in braces cannot exceed $\lambda^k \sum_{i=1}^N p_i^{l+1}$, the conclusion follows directly. □

REMARK. Taking $h(m) := I[m \in A] - \frac{1}{2}$, Corollary 2.4 yields Theorem 1 of Barbour and Hall (1984) when $l = 1$, and an estimate similar to that of the remark following their Theorem 3 when $l = 2$. The sharper inequality (2.17) implies, for $l = 2$, that

$$|\eta_l| \leq 2e_\lambda \sum_{i=1}^N p_i^3 + 2e_\lambda^2\left(\sum_{i=1}^N p_i^2\right)^2,$$

which is almost equivalent to their Theorem 3.

Another interesting example is that of negative binomially distributed summands. Suppose that, for each i,

$$P[X_i = m] = (1 - p_i)^{k_i} p_i^m \binom{k_i + m - 1}{m}, \qquad m \geq 0.$$

Then each X_i can itself be written as a sum of independent components, $X_i = \sum_{r=1}^R X_{ir}$, where

$$P[X_{ir} = m] = (1 - p_i)^{R^{-1}k_i} p_i^m \binom{R^{-1}k_i + m - 1}{m}, \qquad m \geq 0.$$

Now

$$m_{[j]}^{(ir)} = j!\left(\frac{p_i}{1-p_i}\right)^j \binom{R^{-1}k_i + j - 1}{j}$$

and

$$\kappa_{[j]}^{(ir)} = (j-1)!R^{-1}k_i\left(\frac{p_i}{1-p_i}\right)^j,$$

giving

$$\phi_j^{(ir)} = R^{-1}k_i\left(\frac{p_i}{1-p_i}\right)^{j+1}\left\{\binom{R^{-1}k_i + j}{j} + \sum_{t=1}^{j}\binom{R^{-1}k_i + t - 1}{t}\right\},$$

and hence

$$\lambda = \sum_{i,r}\mathbb{E}X_{ir} = \sum_{i=1}^{N} k_i p_i (1-p_i)^{-1},$$

$$\phi_j(R) := \sum_{i,r}\phi_j^{(ir)} = \sum_{i=1}^{N} k_i\left(\frac{p_i}{1-p_i}\right)^{j+1}\left\{\binom{R^{-1}k_i + j}{j} + \sum_{t=1}^{j}\binom{R^{-1}k_i + t - 1}{t}\right\}.$$

We are now in a position to prove the following corollary:

COROLLARY 2.5. *If, for each* $1 \le i \le N$,

$$P[X_i = m] = (1-p_i)^{k_i} p_i^m \binom{k_i + m - 1}{m}, \qquad m \ge 0,$$

then (2.8) *holds with*

$$|\eta_l| \le 2^{2l-1} e_\lambda \left(\sum_{i=1}^{N} k_i\left(\frac{p_i}{1-p_i}\right)^{l+1}\right)\|h\|.$$

PROOF. It follows from (2.12) that (2.8) holds with

$$|\eta_l| \le 2^l \sum_{(s)}\left[\prod_{j=1}^{k}\left(\sum_{i=1}^{N} k_i\left(\frac{p_i}{1-p_i}\right)^{s_j+1}\right)\right] e_\lambda^{k+1}\phi_{s_{k+1}}(R)\|h\|,$$

for each $R \ge 1$, and hence also

$$|\eta_l| \le 2^l \sum_{(s)}\left[\prod_{j=1}^{k}\left(\sum_{i=1}^{N} k_i\left(\frac{p_i}{1-p_i}\right)^{s_j+1}\right)\right] e_\lambda^{k+1}\phi_{s_{k+1}}\|h\|,$$

where

$$\phi_{s_{k+1}} := \lim_{R\to\infty}\phi_{s_{k+1}}(R) = \sum_{i=1}^{N} k_i\left(\frac{p_i}{1-p_i}\right)^{s_{k+1}+1}.$$

The argument is now completed as in Corollary 2.4. □

REMARKS. 1. Taking $N = 1$, Corollary 2.5 yields the estimate

$$d(NB(k, 1-p), P_\lambda) \leq (1 - e^{-\lambda})p(1-p)^{-1},$$

where $\lambda = kp(1-p)^{-1}$ and d denotes total variation distance. This is better than the estimate $\min(2.7, \lambda)p(1-p)^{-1}$ obtained from Kerstan's inequality (2.15). However, the inequalities of Vervaat (1969), with Romanowska's (1977) improvement, provide better estimates of the distance between $NB(k, 1-p)$ and P_λ, except for small λ. This is not too surprising, since they are developed from exact formulae for the distributions being compared, and not merely from the values of a few moments of the summands.

2. Let $k_i = 1$ for all i, so that the X_i are geometrically distributed, with parameters p_i. From Corollary 2.5, the total variation distance from P_λ is at most $e_\lambda \Sigma_{i=1}^N \{p_i/(1-p_i)\}^2$, where $\lambda = \Sigma_{i=1}^N p_i/(1-p_i)$. An alternative way of approximating the sum of the X_i's by a Poisson variate would be to define

$$\tilde{X}_i = \begin{cases} X_i, & X_i = 0, \\ 1, & X_i \geq 1, \end{cases}$$

so that $d(\Sigma_{i=1}^N X_i, \Sigma_{i=1}^N \tilde{X}_i) \leq \Sigma_{i=1}^N p_i^2$, and then to use Corollary 2.4 to show that

$$d\left(\sum_{i=1}^N \tilde{X}_i, P_\mu \right) \leq e_\mu \sum_{i=1}^N p_i^2,$$

where $\mu = \Sigma_{i=1}^N p_i$. To compare the error estimates with the two approaches, consider the case when $p_i = p$, $1 \leq i \leq N$; from the first,

$$d\left(\sum_{i=1}^N X_i, P_\lambda \right) \leq (1 - e^{-\lambda})p(1-p)^{-1},$$

whilst, from the second,

$$d\left(\sum_{i=1}^N X_i, P_\mu \right) \leq (1 - e^{-\mu})p + Np^2.$$

The second of the two estimates is only smaller than both unity and the first of the estimates in the (approximate) ranges $\{N = 1, 0.35 < p < 0.77\}$ and $\{N = 2, 0.45 < p < 0.55\}$.

3. Expansions for expectations of unbounded functions. The expansion (2.8) of Theorem 1 can be extended to cover unbounded functions h, at the cost of some extra effort, largely devoted to establishing a suitable analogue of Corollary 2.3. The following lemma states some necessary estimates:

LEMMA 3.1. (i) *The following upper estimates of*

$$\lambda^{s-1}(m+j)!/(m+j+s-1)!$$

are valid in the ranges indicated:

$$\exp\left\{-\frac{1}{4\lambda}(s-1)(s-2)\right\}, \qquad 1 \le s \le \lambda+2, \quad [\lambda] \le m \le \lambda+\lambda^{1/2}, \quad j \ge 0;$$

$$\exp\{-\tfrac{1}{4}(\lambda-1)\}2^{-(s-2-[\lambda])}, \qquad s > \lambda+2, \quad [\lambda] \le m \le \lambda+\lambda^{1/2}, \quad j \ge 0;$$

$$\exp\left\{-\frac{1}{2\lambda}(s-1)(m-\lambda)\right\}, \qquad s \ge 1, \quad \lambda+\lambda^{1/2} < m \le 2\lambda, \quad j \ge 0;$$

$$2^{-(s-1)}, \qquad s \ge 1, \quad 2\lambda < m, \quad j \ge 0.$$

(ii) *For all* $s \ge 1$, $0 \le m \le [\lambda]$ *and* $0 \le j \le l$,

$$\frac{\lambda^{-s}(m+j)!}{(m+j-s)!} \le \exp\left\{-\frac{s}{2\lambda}(s-1+2(\lambda-m-l))\right\}.$$

PROOF. In case (i), observe that, in $j \ge 0$,

$$\frac{\lambda^{s-1}(m+j)!}{(m+j+s-1)!} \le \prod_{r=1}^{s-1}\left\{1+\frac{m+1-\lambda+r-1}{\lambda}\right\}^{-1}.$$

The estimates now follow from the inequalities

$$(1+x)^{-1} \le \begin{cases} e^{-x/2}, & 0 \le x \le 1, \\ \frac{1}{2}, & x > 1. \end{cases}$$

The proof of case (ii) is similar. □

The next lemma is analogous to Lemma 2.2.

LEMMA 3.2. *Suppose* $l \in \mathbb{N}$ *and* $p > 0$.

(i) *Let* h: $\mathbb{Z}^+ \to \mathbb{R}$ *satisfy* $|\Delta^l h(m)| \le H\{1+m^p\}$. *Then, for all* $\lambda < 1$, $|\Delta^l \theta_\lambda h(m)| \le c_1 H \min(1, m^{-1})\{1+m^p\}$ *for a universal constant* $c_1 = c_1(l, p)$.

(ii) *Let* h: $\mathbb{Z}^+ \to \mathbb{R}$ *satisfy* $|\Delta^l h(m)| \le H\{1+\lambda^{-p/2}|m-[\lambda]|^p\}$, *for some* $\lambda \ge 1$. *Then*

$$|\Delta^l \theta_\lambda h(m)| \le c_2 H \lambda^{-1/2} \min\left(1, \frac{\lambda^{1/2}}{|m-[\lambda]|}\right)\{1+\lambda^{-p/2}|m-[\lambda]|^p\},$$

for a universal constant $c_2 = c_2(l, p)$.

PROOF. Observe first that

$$\theta_\lambda C_n(\lambda; m) = -\frac{1}{\lambda} C_{n-1}(\lambda; m), \qquad n \ge 1, \quad m \ge 0,$$

and that

$$\theta_\lambda C_0(\lambda; m) = 0.$$

Hence, for any polynomial π of degree $l-1$, $\theta_\lambda \pi$ is a polynomial of degree $l-2$, so that

$$\Delta^l(h+\pi) = \Delta^l h \quad \text{and} \quad \Delta^l \theta_\lambda (h+\pi) = \Delta^l \theta_\lambda h.$$

Thus we may without loss of generality assume that

$$\Delta^r h([\lambda]) = 0, \qquad 0 \le r \le l-1. \tag{3.1}$$

Next, use Equation (2.3) to express $\theta_\lambda h$ as

$$\theta_\lambda h(m) = \sum_{s=0}^{m} \lambda^{-1-s} \frac{m!}{(m-s)!}\big(h(m-s) - \bar{h}\big) \tag{3.2}$$

$$= -\sum_{s\ge 1} \frac{m!}{(m+s)!}\lambda^{s-1}\big(h(m+s) - \bar{h}\big), \tag{3.3}$$

so that

$$\Delta^l \theta_\lambda h(m) = \sum_{s=0}^{m} \lambda^{-1-s} \sum_{j=0\vee(l-s)}^{l} \binom{l}{j} \big\{\Delta^j h(m-s) - \bar{h}\delta_{j0}\big\} \times \frac{s!(m+j)!}{(s-l+j)!(m+l-s)!} \tag{3.4}$$

$$= -\sum_{s\ge 1} \lambda^{s-1} \sum_{j=0}^{l} \binom{l}{j} \big\{\Delta^j h(m+s) - \bar{h}\delta_{j0}\big\}(-1)^{l-j} \times \frac{(s+l-j-1)!(m+j)!}{(s-1)!(m+l+s)!}. \tag{3.5}$$

For case (i), it follows from (3.1) and the bound on $\Delta^l h$ that

$$|\Delta^j h(m)| \le \frac{m^{l-j}}{(l-j)!} H\{1 + m^p\}, \qquad 0 \le j \le l, \tag{3.6}$$

and so, in particular,

$$|\bar{h}| \le H\{\mathbb{E}Q^l + \mathbb{E}Q^{l+p}\}/l! \le k_0 H,$$

uniformly in $\lambda \le 1$, where, here and subsequently, k_j denotes a universal constant, depending only on l and p. Hence the contribution from $\bar{h}$ to the expression (3.5) is no larger than

$$k_0 H \sum_{s\ge 1} \frac{(s+l-1)!m!}{(s-1)!(m+l+s)!} \le k_0 H \sum_{s\ge 1} \frac{1}{(s-1)!(m+l+s)} \le k_1 H (m+1)^{-1}. \tag{3.7}$$

The conclusion in case (i) will now follow from (3.5), if it can be shown that

$$\sum_{s\ge 1} |\Delta^j h(m+s)| \frac{(s+l-j-1)!(m+j)!}{(s-1)!(m+l+s)!} \le k_2 (m+1)^{-1} H\{1+m^p\}, \tag{3.8}$$

for $0 \le j \le l$. However,

$$\frac{(s-1+l-j)!}{(s-1)!}\frac{(m+j)!}{(m+l+s)!} \le \frac{(s-1+l-j)^{l-j}}{(m+j+s)^{l-j}}\frac{1}{(s-1)!(m+l+s)},$$

and hence, from (3.6), the left-hand side of (3.8) is majorized by

$$k_3 H \sum_{s\ge 1} \frac{s^l(s^p+m^p)}{(s-1)!(m+s+1)} \le \frac{k_2 H\{1+m^p\}}{(m+1)},$$

as required.

For case (ii), we have

$$|\Delta^j h(m)| \le \frac{|m-[\lambda]|^{l-j}}{(l-j)!} H\{1+\lambda^{-p/2}|m-[\lambda]|^p\}, \qquad 0 \le j \le l, \tag{3.9}$$

in place of (3.6), and, in particular,

$$|\bar{h}| \le Hk_4\lambda^{l/2}. \tag{3.10}$$

Taking first the case $m \ge [\lambda]$, and using (3.5) and (3.9), it is necessary to bound

$$\sum_{s\ge 1} \lambda^{s-1}(m+s-[\lambda])^{l-j}\{1+\lambda^{-p/2}(m+s-[\lambda])^p\} \times \frac{(s+l-j-1)!(m+j)!}{(s-1)!(m+l+s)!}, \qquad 0 \le j \le l, \tag{3.11}$$

and, from (3.10),

$$\sum_{s\ge 1} \lambda^{s-1+l/2} \frac{(s+l-1)!m!}{(s-1)!(m+l+s)!}. \tag{3.12}$$

Now the expression in (3.11) is no greater than

$$2^{2(l-j)+p} \sum_{s\ge 1} \left\{\frac{\lambda^{s-1}(m+j)!}{(m+j+s-1)!}\right\} \frac{\{(s-1)^{l-j}+(l-j)^{l-j}\}}{m^{l+1-j}} \times \{(m-[\lambda])^{l-j}+s^{l-j}\}\{1+\lambda^{-p/2}(m-[\lambda])^p+\lambda^{-p/2}s^p\}, \tag{3.13}$$

which can be estimated using Lemma 3.1(i). For instance, it follows from Lemma 3.1(i) that, for $[\lambda] \le m \le \lambda + \sqrt{\lambda}$, the essential contribution to (3.13) comes from terms with $s = O(\lambda^{1/2})$, from which it follows easily that, for such m, (3.13) is of order

$$\lambda^{1/2} \cdot \lambda^{l-j} m^{-(l-j+1)} = O(\lambda^{-1/2});$$

for $\lambda + \sqrt{\lambda} < m \le 2\lambda$, the terms with $s = O(\lambda/(m-\lambda))$ are important, and (3.13) has order

$$\frac{\lambda}{(m-\lambda)}\left(\frac{\lambda}{m-\lambda}\right)^{l-j}\frac{(m-\lambda)^{l-j}}{m^{l+1-j}}\{1+(m-[\lambda])^p\lambda^{-p/2}\} \le (m-\lambda)^{-1}\{1+(m-[\lambda])^p\lambda^{-p/2}\};$$

for $m > 2\lambda$, the terms with $s = O(1)$ are the essential ones, and (3.13) has order

$$m^{-1}\{1 + (m - [\lambda])^p \lambda^{-p/2}\}.$$

The estimation of (3.12) is accomplished by a similar argument, and it follows that, for $m \geq [\lambda]$,

$$(3.14)\quad |\Delta^l \theta_\lambda h(m)| \leq k_5 H \lambda^{-1/2} \min\left(1, \frac{\lambda^{1/2}}{|m - [\lambda]|}\right)\{1 + \lambda^{-p/2}|m - [\lambda]|^p\}.$$

For $m < [\lambda]$, use (3.4) and argue similarly. The principal quantities to be estimated are no greater than

$$\sum_{s=0}^{m} \lambda^{-1}\left\{\frac{\lambda^{-s}(m+j)!}{(m+j-s)!}\right\}\left(\frac{s}{m+1-s}\right)^{l-j}\{s^{l-j} + (\lambda - m)^{l-j}\}$$
$$\times\{1 + \lambda^{-p/2}s^p + \lambda^{-p/2}|m - [\lambda]|^p\},$$

and Lemma 3.1 (ii) is used to conclude that (3.14) holds also for $m < [\lambda]$. □

COROLLARY 3.3. *Suppose $p > 0$ and $l, r, s_1 \ldots, s_r \in \mathbb{N}$ are such that $\sum_{j=1}^r s_j = l$. Then there exist universal constants $c_3(l, p)$ and $c_4(l, p)$ such that:*

(i) *if $h\colon \mathbb{Z}^+ \to \mathbb{R}$ satisfies $|\Delta^l h(m)| \leq H\{1 + m^p\}$, then*

$$\left|\prod_{j=1}^{r} (\Delta^{s_j}\theta_\lambda)h(m)\right| \leq c_3 H \min(1, m^{-1})\{1 + m^{p-r+1}\},$$

for all $\lambda < 1$;

(ii) *if $h\colon \mathbb{Z}^+ \to \mathbb{R}$ satisfies $|\Delta^l h(m)| \leq H\{1 + \lambda^{-p/2}|m - [\lambda]|^p\}$ for some $\lambda \geq 1$, then*

$$\left|\prod_{j=1}^{r} (\Delta^{s_j}\theta_\lambda)h(m)\right| \leq c_4 H \lambda^{-r/2} \min\left(1, \frac{\lambda^{1/2}}{|m - [\lambda]|}\right)$$
$$\times\left\{1 + \left(\lambda^{-1/2}|m - [\lambda]|\right)^{p-r+1}\right\}.$$

PROOF. Immediate, from Lemma 3.2 □

Having established a counterpart to Corollary 2.3, it is necessary to find an analogue of Lemma 2.1. The following lemma is a preliminary step.

LEMMA 3.4. (i) *Suppose $\lambda \geq 1$ and $\mathbb{E}X_i^2 < \infty$, $1 \leq i \leq N$. Then*

$$\mathbb{E}\{\lambda^{-p/2}|W - \lambda|^p\} \leq \left(1 + \lambda^{-1}\nu_1\right)^{p/2}, \quad \text{for } 0 \leq p \leq 2.$$

If, for some $p > 2$, $\mathbb{E}X_i^p < \infty$, $1 \leq i \leq N$,

$$\mathbb{E}\{\lambda^{-p/2}|W - \lambda|^p\} \leq c_5\left\{1 + \lambda^{-p/2}\sum_{i=1}^{N} \mathbb{E}X_i^p + \left(\lambda^{-1}\nu_1\right)^{p/2}\right\},$$

for a universal constant $c_5(p)$.

(ii) *If* $\lambda \le 1$, *the inequalities*

$$\mathbb{E}(W^p) \le (2 + \nu_1)^{p/2}, \qquad 0 \le p \le 2,$$

and

$$\mathbb{E}(W^p) \le c_6\left\{1 + \sum_{i=1}^{N} \mathbb{E}X_i^p + \nu_1^{p/2}\right\}, \qquad p > 2,$$

are valid.

PROOF. By direct computation, in the case $\lambda \ge 1$,

$$\mathbb{E}(W - \lambda)^2 = \sum_{i=1}^{N} \operatorname{var} X_i = \lambda + \sum_{i=1}^{N} \left(m_{[2]}^{(i)} - m_{[1]}^{(i)^2}\right) \le \lambda + \nu_1,$$

and the inequality for $p \le 2$ follows from Hölder's inequality. For $p > 2$, use an inequality of Marcinkiewicz and Zygmund (1937) form: For $p > 2$,

$$\mathbb{E}|W - \lambda|^p \le C_p\left\{\sum_{i=1}^{N} \mathbb{E}|X_i - \mathbb{E}X_i|^p + \left(\sum_{i=1}^{N} \operatorname{var} X_i\right)^{p/2}\right\},$$

for a universal constant C_p, from which the result now follows. The proof for $\lambda \le 1$ is the same. □

We now prove the following counterpart to Lemma 2.1:

LEMMA 3.5. *For some* $l \in \mathbb{N}$ *and* $q \ge 0$, *suppose that* $\mathbb{E}X_i^\alpha < \infty$, $1 \le i \le N$, *for* $\alpha = 1 + \max(q, 1)$. *Suppose also that* $\lambda \ge 1$, *and that* $\nu_1 \le \lambda$, *and let* $g\colon \mathbb{Z}^+ \to \mathbb{R}$ *satisfy*

$$|\Delta^l g(m)| \le G \min\left(1, \lambda^{1/2}|m - [\lambda]|^{-1}\right)\left\{1 + \lambda^{-q/2}|m - [\lambda]|^q\right\},$$

for some $G < \infty$. *Then*

$$e_l(g) := \mathbb{E}\{Wg(W) - \lambda g(W+1)\} - \sum_{s=1}^{l-1} \frac{1}{s!}\kappa_{[s+1]}\mathbb{E}\{\Delta^s g(W+1)\}$$

satisfies

$$|e_l(g)| \le c_7 G\left\{\nu_l + \lambda^{-(q-1)/2}\nu_{l,q}\right\},$$

for a universal constant $c_7 = c_7(l, q)$, *where*

$$\nu_{l,q} := \begin{cases} \sum_{i=1}^{N} \mathbb{E}\{\psi_{l,q}(X_i)\}, & q > 1, \\ 0, & q \le 1. \end{cases}$$

If $\lambda \le 1$ *and* $\nu_1 \le 1$, *and* $g\colon \mathbb{Z}^+ \to \mathbb{R}$ *satisfies*

$$|\Delta^l g(m)| \le G \min(1, m^{-1})\{1 + m^q\},$$

then

$$|e_l(g)| \le c_8 G(\nu_1 + \nu_{1,q}).$$

REMARK. If $\nu_1 \geq \lambda$, no useful Poisson approximation can be expected.

PROOF. We prove only the case $\lambda \geq 1$. The other proof is similar. Under the given conditions on g, all the required expectations exist.

For the case $q > 1$, proceed as for Lemma 2.1, using inequality (1.3) of Lemma 1.1 and the given bound for $|\Delta^l g(m)|$, obtaining easily that the quantity to be estimated is no larger than

$$k_6 G \sum_{i=1}^{N} \mathbb{E}\Big\{\Big(1 + \mathbb{E}\big|\lambda^{-1/2}(W_i - [\lambda])\big|^{q-1} + |\lambda^{-1/2}X_i|^{q-1}\Big)\psi_{l,1}(X_i)\Big\},$$

where

$$\mathbb{E}\big|\lambda^{-1/2}(W_i - [\lambda])\big|^{q-1} \leq 3^{q-1}\Big\{1 + \mathbb{E}\big|\lambda^{-1/2}(W - \lambda)\big|^{q-1} + \mathbb{E}|\lambda^{-1/2}X_i|^{q-1}\Big\}.$$

Now apply Lemma 3.4. For $0 \leq q \leq 1$, $\Delta^l g$ is bounded, and Lemma 2.1 can be directly applied. □

The theorem can now be stated.

Define

$$\mu_{l,0}^{\lambda} := \max_{(s)}\left\{\frac{\lambda^{-(k+1)/2}\nu_{s_{k+1}}}{(s_{k+1}-1)!}\left|\prod_{j=1}^{k}\frac{\kappa_{[s_j+1]}}{s_j!}\right|\right\},$$

$$\mu_{l,p}^{\lambda} := \max_{(s)}\left\{\frac{\lambda^{-p/2}\nu_{s_{k+1},p-k}}{(s_{k+1}-1)!}\left|\prod_{j=1}^{k}\frac{\kappa_{[s_j+1]}}{s_j!}\right|\right\},$$

where, as before, $\max_{(s)}$ is taken over

$$\left\{k \geq 0;\ s_j \geq 1, 1 \leq j \leq k+1;\ \sum_{j=1}^{k+1} s_j = l\right\}.$$

THEOREM 2. *Suppose $l \in \mathbb{N}$, $p > 0$ and $\mathbb{E}X_i^{\alpha} < \infty$, $1 \leq i \leq N$, where $\alpha = l + \max(1, p)$. Suppose also that $\lambda \geq 1$ and that $\nu_1 \leq \lambda$. Let h: $\mathbb{Z}^+ \to \mathbb{R}$ satisfy $|\Delta^l h(m)| \leq H\{1 + \lambda^{-p/2}|m - [\lambda]|^p\}$. Then*

$$\mathbb{E}h(W) = \int h\,dQ_l + \eta_l, \tag{3.15}$$

where

$$|\eta_l| \leq cH\{\mu_{l,0}^{\lambda} + \mu_{l,p}^{\lambda}\}, \tag{3.16}$$

and $c = c(l, p)$ is a universal constant.

If $\lambda \leq 1$ and $\nu_1 \leq 1$, and h: $\mathbb{Z}^+ \to \mathbb{R}$ satisfies $|\Delta^l h(m)| \leq H\{1 + m^p\}$, then (3.15) *holds with*

$$|\eta_l| \leq cH\{\mu_{l,0}^{1} + \mu_{l,p}^{1}\}. \tag{3.17}$$

PROOF. As in the proof of Theorem 1,

$$\mathbb{E}h(W) = \sum_{(s)}(-1)^k\left(\prod_{j=1}^{k}\frac{\kappa_{[s_j+1]}}{s_j!}\right)\mathbb{E}\left\{\prod_{j=1}^{k}(\Delta^{s_j}\theta_\lambda)h(Q)\right\} + \eta_l',$$

where

$$|\eta_l'| \le \sum_{(s)}\left|\prod_{j=1}^{k}\frac{\kappa_{[s_j+1]}}{s_j!}\right|e_{s_{k+1}}\left(\theta_\lambda\prod_{j=1}^{k}(\Delta^{s_j}\theta_\lambda)h\right).$$

As before, $\Sigma_{(s)}$ denotes the sum over

$$\left\{k \ge 0;\ s_j \ge 1, 1 \le j \le k+1;\ \sum_{j=1}^{k+1} s_j = l\right\}.$$

Using Corollary 3.3 and Lemma 3.5, it thus follows that

$$\begin{aligned}|\eta_l'| &\le \sum_{(s)}\left|\prod_{j=1}^{k}\frac{\kappa_{[s_j+1]}}{s_j!}\right|c_4(l,p)H\lambda^{-(k+1)/2}c_7\left\{\nu_{s_{k+1}} + \lambda^{-(p-k-1)/2}\nu_{s_{k+1},p-k}\right\}\\ &\le cH\left\{\mu_{l,0}^{\lambda} + \mu_{l,p}^{\lambda}\right\},\end{aligned}$$

from the definition of μ^λ. □

REMARKS. 1. For $p \le 1$, $\mu_{l,p}^{\lambda} = 0$, and the error estimate is of a similar form to that of Theorem 1. Note, however, that, in the definition of $\mu_{l,0}^{\lambda}$, there enters a factor $\lambda^{-(k+1)/2}$ in place of the factor e_λ^{k+1} in the definition of μ_l, which, for large λ, is of the smaller order $\lambda^{-(k+1)}$.

2. Suppose that, for some $p \ge 1$, $\mathbb{E}X_i^{p+1} < \infty$, $1 \le i \le N$, and $\nu_1 \le \lambda \ge 1$. Let $A = \{j\colon P[W=j] \ge P[Q=j]\}$, and take

$$h(j) := \lambda^{-p/2}|j - [\lambda]|^p\{2I[j \in A] - 1\}.$$

Then, using Theorem 2 with $l = 1$,

$$0 \le \mathbb{E}h(W) - \mathbb{E}h(Q) \le 2c\left\{\lambda^{-1/2}\nu_1 + \lambda^{-p/2}\nu_{1,p}\right\}.$$

Hence the total variation distance between the distribution of W and P_λ, restricted to the set $\{j\colon \lambda^{-1/2}|j - [\lambda]| \ge m\}$, is no larger than a constant times $m^{-p}\{\lambda^{-1/2}\nu_1 + \lambda^{-p/2}\nu_{1,p}\}$.

3. In the case of 0–1 summands, ν_1 is automatically less than λ, and

$$\begin{aligned}\mu_{l,p}^{\lambda} \le \mu_{l,0}^{\lambda} &= \max_{(s)}\left\{\lambda^{-(k+1)/2}\prod_{j=1}^{k+1}\left(\sum_{i=1}^{N}p_i^{s_j+1}\right)\right\}\\ &\le \lambda^{l/2-1}\sum_{i=1}^{N}p_i^{l+1}.\end{aligned}$$

Thus, for example, in the binomial case, useful expansions for the expectations of at most polynomially growing functions of $B(n,p)$ are only obtained from Theorem 2 when np^3 is small.

REFERENCES

BARBOUR, A. D. (1986). Asymptotic expansions based on smooth functions in the central limit theorem. *Probab. Theory Rel. Fields* **72** 289–303.

BARBOUR, A. D. and EAGLESON, G. K. (1983). Poisson approximation for some statistics based on exchangeable trials. *Adv. in Appl. Probab.* **15** 585–600.

BARBOUR, A. D. and HALL, P. (1984). On the rate of Poisson convergence. *Math. Proc. Cambridge Philos. Soc.* **95** 473–480.

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

DEHEUVELS, P. and PFEIFER, D. (1986). A semigroup approach to Poisson approximation. *Ann. Probab.* **14** 665–678.

KENDALL, M. G. and STUART, A. (1963). *The Advanced Theory of Statistics* **1**. Griffin, London.

KERSTAN, J. (1964). Verallgemeinerung eines Satzes von Procharov und Le Cam. *Z. Wahrsch. verw. Gebiete* **2** 173–179.

MARCINKIEWICZ, J. and ZYGMUND, A. (1937). Sur les fonctions indépendantes. *Fund. Math.* **29** 60–90.

ROMANOWSKA, M. (1977). A note on the upper bound for the distance in total variation between the binomial and the Poisson distribution. *Statist. Neerlandica* **31** 127–130.

SERFLING, R. J. (1975). A general Poisson approximation theorem. *Ann. Probab.* **3** 726–731.

VERVAAT, W. (1969). Upper bounds for the distance in total variation between the binomial or negative binomial and the Poisson distribution. *Statist. Neerlandica* **23** 79–86.

INSTITUT FÜR ANGEWANDTE MATHEMATIK
UNIVERSITÄT ZÜRICH
RÄMISTRASSE 74
CH-8001 ZÜRICH
SWITZERLAND