Bernoulli **6**(4), 2000, 581–590

Estimating Stein's constants for compound Poisson approximation

A.D. BARBOUR[1] and AIHUA XIA[2]

[1]*Institut für Angewandte Mathematik, Universität Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland. E-mail: adb@amath.unizh.ch*
[2]*Department of Statistics, School of Mathematics, The University of New South Wales, Sydney 2052, Australia. E-mail: xia@maths.unsw.edu.au*

Stein's method for compound Poisson approximation was introduced by Barbour, Chen and Loh. One difficulty in applying the method is that the bounds on the solutions of the Stein equation are by no means as good as for Poisson approximation. We show that, for the Kolmogorov metric and under a condition on the parameters of the approximating compound Poisson distribution, bounds comparable with those obtained for the Poisson distribution can be recovered.

Keywords: coupling; immigration–death process; Kolmogorov metric; Stein's method

1. Introduction

Let W be any random variable on $\mathbb{Z}_+$, and let $\lambda_i \geqslant 0$, $i \in \mathbb{N}$, be chosen to satisfy $\sum_{i\geqslant 1} i\lambda_i < \infty$. Suppose that it can be shown that

$$\left|\mathrm{E}\left\{\sum_{i\geqslant 1} i\lambda_i g(W+i) - Wg(W)\right\}\right| \leqslant \varepsilon_0 M_0(g) + \varepsilon_1 M_1(g) \tag{1.1}$$

for all bounded g: $\mathbb{N} \to \mathbb{R}$, where $M_l(g) := \sup_{w\in\mathbb{N}} |\Delta^l g(w)|$, $l \in \mathbb{Z}_+$, and $\Delta g(w) := g(w+1) - g(w)$. Then it follows that

$$d_{\mathscr{F}}(\mathscr{L}(W), \mathrm{CP}(\boldsymbol{\lambda})) := \sup_{f\in\mathscr{F}} |\mathrm{E}f(W) - \mathrm{CP}(\boldsymbol{\lambda})\{f\}| \leqslant \varepsilon_0 \sup_{f\in\mathscr{F}} M_0(g_f) + \varepsilon_1 \sup_{f\in\mathscr{F}} M_1(g_f), \tag{1.2}$$

for any set $\mathscr{F}$ of test functions, where g_f solves the Stein equation

$$\sum_{i\geqslant 1} i\lambda_i g(j+i) - jg(j) = f(j) - \mathrm{CP}(\boldsymbol{\lambda})\{f\}, \qquad j \geqslant 0. \tag{1.3}$$

Here, $\mathrm{CP}(\boldsymbol{\lambda})$ denotes the compound Poisson distribution of $\sum_{i\geqslant 1} iZ_i$, where the $Z_i \sim \mathrm{Po}(\lambda_i)$ are independent.

There are many occasions, some of them discussed in Roos (1994), in which (1.1) can be shown to hold for small ε_0 and ε_1. However, the resulting distance estimates (1.2) are not as powerful as they could be, for lack of sharp bounds on the quantities $\sup_{f\in\mathscr{F}} M_l(g_f)$ for the commonest choices of test functions $\mathscr{F}$ and for most $\mathrm{CP}(\boldsymbol{\lambda})$. Barbour *et al.* (1992a) found

1350–7265

reasonable bounds for the test functions $\mathscr{F}_{\mathrm{TV}} = \{1_A, A \subset \mathbb{Z}_+\}$, appropriate to total variation approximation, under the additional condition on the λ_i that

$$\lambda_1 \geqslant 2\lambda_2 \geqslant 3\lambda_3 \geqslant \ldots; \tag{1.4}$$

their bounds are

$$\sup_{f \in \mathscr{F}_{\mathrm{TV}}} M_0(g_f) \leqslant \left(1 \wedge \frac{2}{\sqrt{\rho}}\right); \qquad \sup_{f \in \mathscr{F}_{\mathrm{TV}}} M_1(g_f) \leqslant \left\{1 \wedge \frac{1}{\rho}\left(\frac{1}{4\rho} + \log^+(2\rho)\right)\right\}, \tag{1.5}$$

where $\rho = \lambda_1 - 2\lambda_2$. The bound on M_1 is weak because of the logarithmic factor, which may be superfluous. In this paper, we consider only the set of test functions $\mathscr{F}_{\mathrm{K}} := \{f_k,\ k \in \mathbb{N}\colon f_k(x) = 1_{[k,\infty)}(x)\}$ appropriate to Kolmogorov distance. For these functions, we give neat bounds which do not involve any logarithmic factor, and which replace $\lambda_1 - 2\lambda_2$ in the denominator by λ_1, at times also a substantial improvement: these are contained in the following result.

Proposition 1.1. *Let g_k denote the solution to the Stein equation (1.3) for $f = f_k$. If condition (1.4) holds, then, for all $k \in \mathbb{N}$,*

$$M_0(g_k) \leqslant 1 \wedge \sqrt{\frac{2}{\mathrm{e}\lambda_1}}, \tag{1.6}$$

$$M_1(g_k) \leqslant \frac{1}{2} \wedge \frac{1}{\lambda_1 + 1}. \tag{1.7}$$

Remark 1.2. Under condition (1.4), our bounds (1.6) and (1.7) are uniformly sharper than those in Theorem 3.1 of Barbour and Utev (1998); in particular, there is no unwanted logarithmic factor in (1.7), nor do our bounds become large if $2\lambda_2$ is close to λ_1.

We prove the proposition by using probabilistic arguments. To introduce them, let $\nu_i = i\lambda_i - (i+1)\lambda_{i+1}$, $i \geqslant 1$. Under condition (1.4), the Stein equation (1.3) can be rephrased in terms of a function h such that $g = \Delta h$, in the form

$$\mathscr{C} h(n) = f(n) - \mathrm{CP}(\boldsymbol{\lambda})(f), \qquad n \in \mathbb{Z}_+, \tag{1.8}$$

where the generator $\mathscr{C}$, defined by

$$\mathscr{C} h(n) = \sum_{i=1}^{\infty} [h(n+i) - h(n)]\nu_i + n[h(n-1) - h(n)], \tag{1.9}$$

is that of an immigration–death process X with unit per capita death rate and with immigration in batches at intensity λ_1, a batch of size j coming with probability ν_j/λ_1. X has equilibrium distribution $\mathrm{CP}(\boldsymbol{\lambda})$, and the Stein equation (1.8) has solution h_f given by

$$h_f(n) = -\int_0^{\infty} [\mathrm{E} f(X_n(t)) - \mathrm{CP}(\boldsymbol{\lambda})(f)]\, \mathrm{d}t, \tag{1.10}$$

where X_n is an X-process with $X_n(0) = n$. Note that X_n, X_{n+1} and X_{n+2} can be realized on the same probability space by taking E_1 and E_2 to be two independent standard exponential random variables which are also independent of X_n, and setting

$$X_{n+1}(t) = X_n(t) + 1_{\{E_1 > t\}}, \qquad X_{n+2}(t) = X_n(t) + 1_{\{E_1 > t\}} + 1_{\{E_2 > t\}}. \tag{1.11}$$

Let h_k denote the solution to the Stein equation (1.10) for $f = f_k$, so that we have $g_k = \Delta h_k$. Then it follows that

$$\begin{aligned}\delta_1 h_k(n) &:= -[h_k(n+1) - h_k(n)] = -g_k(n),\\ \delta_2 h_k(n) &:= -[h_k(n+2) - 2h_k(n+1) + h_k(n)] = -\Delta g_k(n),\end{aligned}$$

and the required bounds on $M_0(g_k)$ and $M_1(g_k)$ follow from corresponding bounds on $\delta_1 h_k(n)$ and $\delta_2 h_k(n)$, $n \in \mathbb{Z}_+$. Now (1.11) and (1.10) can immediately be used to give

$$\begin{aligned}\delta_1 h_k(n) &= \int_0^\infty \mathrm{E}[f_k(X_n(t) + 1_{\{E_1 > t\}}) - f_k(X_n(t))]\,\mathrm{d}t\\ &= \int_0^\infty \mathrm{e}^{-t}\mathrm{E}[f_k(X_n(t) + 1) - f_k(X_n(t))]\,\mathrm{d}t, \end{aligned} \tag{1.12}$$

$$\begin{aligned}\delta_2 h_k(n) &= \int_0^\infty \mathrm{E}[f_k(X_n(t) + 1_{\{E_1 > t\}} + 1_{\{E_2 > t\}}) - f_k(X_n(t) + 1_{\{E_1 > t\}})\\ &\qquad - f_k(X_n(t) + 1_{\{E_2 > t\}}) + f_k(X_n(t))]\,\mathrm{d}t\\ &= \int_0^\infty \mathrm{e}^{-2t}\,\mathrm{E}[f_k(X_n(t) + 2) - 2f_k(X_n(t) + 1) + f_k(X_n(t))]\,\mathrm{d}t. \end{aligned} \tag{1.13}$$

Clearly,

$$f_k(X_n(t) + 1) - f_k(X_n(t)) = \begin{cases} 1, & \text{if } X_n(t) = k - 1,\\ 0, & \text{otherwise,}\end{cases} \tag{1.14}$$

and

$$f_k(X_n(t) + 2) - 2f_k(X_n(t) + 1) + f_k(X_n(t)) = \begin{cases} 0, & \text{if } X_n(t) \geqslant k \text{ or } \leqslant k - 3,\\ -1, & \text{if } X_n(t) = k - 1,\\ 1, & \text{if } X_n(t) = k - 2.\end{cases} \tag{1.15}$$

The combination of the representations (1.12) and (1.13) with the very simple forms of the integrands given in (1.14) and (1.15) makes the proofs possible. Indeed, it already follows immediately that $\delta_1 h_k(n) \geqslant 0$ for all n and k, and that $M_1(h_k) \leqslant 1$ and $M_2(h_k) \leqslant \frac{1}{2}$.

2. Proof of (1.6)

For (1.6), we use (1.12), writing X_n in the form $X_n(t) = Y_n(t) + S(t)$, where Y_n and S are independent, S denoting the population resulting from immigrants after time 0, and Y_n that remaining from the initial n individuals at time 0. Then, by the usual concentration inequality,

$$\max_{s\geq 0}\mathrm{P}(X_n(t)=s)\leq \max_{s\geq 0}\mathrm{P}(S(t)=s). \tag{2.1}$$

Fixing any $t>0$, the number of batches immigrating between 0 and t has a Poisson distribution; conditional on this number, the times of immigration are independent, and uniformly distributed on $[0, t]$. Let p_t denote the probability that a batch arriving in $[0, t]$ has individuals still surviving at time t. Then

$$p_t = t^{-1}\int_0^t \sum_{i\geq 1}(\nu_i/\lambda_1)\{1-(1-\mathrm{e}^{-u})^i\}\,\mathrm{d}u \geq t^{-1}\int_0^t \mathrm{e}^{-u}\,\mathrm{d}u = t^{-1}(1-\mathrm{e}^{-t}), \tag{2.2}$$

and hence the number N_t of batches which arrive in $[0, t]$ and have individuals still alive at t, a thinning of the original batches, has distribution $\mathrm{Po}(\lambda_1 t p_t)$.

Let U_l, $l\in\mathbb{N}$, be independent, and distributed according to the number of members of a batch arriving in $[0, t]$ which are alive at time t, *conditional* on there being at least one alive. Then $\mathrm{P}(S(t)=0)=\mathrm{P}(N_t=0)$ and

$$\mathrm{P}(S(t)=s)=\sum_{r\geq 0}\mathrm{P}(N_t=r)\mathrm{P}\left(\sum_{l=1}^r U_l = s\right). \tag{2.3}$$

But, for $s\geq 1$,

$$\sum_{r\geq 1}\mathrm{P}\left(\sum_{l=1}^r U_l = s\right)=\mathrm{P}\left(\bigcup_{r\geq 1}\left\{\sum_{l=1}^r U_l = s\right\}\right)\leq 1,$$

so that, for all $s\geq 0$,

$$\mathrm{P}(S(t)=s)\leq \max_{r\geq 0}\mathrm{P}(N_t=r)\leq \{2\,\mathrm{e}\lambda_1(1-\mathrm{e}^{-t})\}^{-1/2}, \tag{2.4}$$

(see Barbour *et al.* 1992b, p. 262). Combining this with (1.12) and (1.14), it follows that

$$\delta_1 h_k(n)\leq \int_0^\infty \frac{\mathrm{e}^{-t}}{\sqrt{2\,\mathrm{e}\lambda_1(1-\mathrm{e}^{-t})}}\,\mathrm{d}t=\sqrt{\frac{2}{\mathrm{e}\lambda_1}},$$

as required.

3. Proof of (1.7)

We begin with a straightforward calculation. If Z is an exponential random variable with mean $1/\mu$, then

$$\mathrm{E}\exp(-2Z)=\frac{\mu}{\mu+2}, \qquad \mathrm{E}\left\{\int_0^Z \mathrm{e}^{-2t}\,\mathrm{d}t\right\}=\frac{1}{\mu+2}. \tag{3.1}$$

Now, defining $S_j^i=\inf\{t\colon X_i(t)=j\}$, $i, j\geq 0$, we have

$$f_k(X_n(t)+2) - 2f_k(X_n(t)+1) + f_k(X_n(t)) \geqslant 0,$$

for $t < S^n_{k-1}$. Thus, from (1.13), it follows that

$$\begin{aligned}
\delta_2 h_k(n) &= \mathrm{E}\int_0^{S^n_{k-1}} \mathrm{e}^{-2t}[f_k(X_n(t)+2) - 2f_k(X_n(t)+1) + f_k(X_n(t))]\,\mathrm{d}t \\
&\quad + \mathrm{E}\int_{S^n_{k-1}}^{\infty} \mathrm{e}^{-2t}[f_k(X_n(t)+2) - 2f_k(X_n(t)+1) + f_k(X_n(t))]\,\mathrm{d}t \\
&\geqslant \mathrm{E}\int_{S^n_{k-1}}^{\infty} \mathrm{e}^{-2t}[f_k(X_n(t)+2) - 2f_k(X_n(t)+1) + f_k(X_n(t))]\,\mathrm{d}t \\
&= \mathrm{E}\exp(-2S^n_{k-1}) \times \mathrm{E}\int_0^{\infty} \mathrm{e}^{-2t}[f_k(X_{k-1}(t)+2) - 2f_k(X_{k-1}(t)+1) + f_k(X_{k-1}(t))]\,\mathrm{d}t \\
&= \mathrm{E}\exp(-2S^n_{k-1})\delta_2 h_k(k-1). \qquad (3.2)
\end{aligned}$$

Similarly, since $f_k(X_n(t)+2) - 2f_k(X_n(t)+1) + f_k(X_n(t)) \leqslant 0$ for $t < S^n_{k-2}$, we obtain

$$\begin{aligned}
\delta_2 h_k(n) &= \mathrm{E}\int_0^{S^n_{k-2}} \mathrm{e}^{-2t}[f_k(X_n(t)+2) - 2f_k(X_n(t)+1) + f_k(X_n(t))]\,\mathrm{d}t \\
&\quad + \mathrm{E}\int_{S^n_{k-2}}^{\infty} \mathrm{e}^{-2t}[f_k(X_n(t)+2) - 2f_k(X_n(t)+1) + f_k(X_n(t))]\,\mathrm{d}t \\
&\leqslant \mathrm{E}\exp(-2S^n_{k-2})\delta_2 h_k(k-2). \qquad (3.3)
\end{aligned}$$

Thus, in order to bound $\delta_2 h_k(n)$, it is enough to be able to control $\delta_2 h_k(k-1)$ and $\delta_2 h_k(k-2)$.

Next, we show that

$$\delta_2 h_k(k-1) \leqslant 0, \qquad \text{for } k \geqslant 1 \qquad (3.4)$$

$$\delta_2 h_k(k-2) > 0, \qquad \text{for } k > 1. \qquad (3.5)$$

First, observe that, for $r \geqslant k-1$,

$$\begin{aligned}
\sum_{n=k-1}^{r} \delta_2 h_k(n) &= \sum_{n=k-1}^{r} [\delta_1 h_k(n+1) - \delta_1 h_k(n)] \\
&= \delta_1 h_k(r+1) - \delta_1 h_k(k-1),
\end{aligned}$$

and that

$$\delta_1 h_k(r+1) = \int_0^{\infty} \mathrm{e}^{-t}\,\mathrm{E}[f_k(X_{r+1}(t)+1) - f_k(X_{r+1}(t))]\,\mathrm{d}t = \mathrm{E}\exp\{-S^{r+1}_{k-1}\}\delta_1 h_k(k-1) \to 0 \quad \text{as } r \to \infty,$$

since $\lim_{r\to\infty} S^r_{k-1} = \infty$ almost surely. Hence $\sum_{n=k-1}^{\infty}\delta_2 h_k(n)$ converges, and

$$\sum_{n=k-1}^{\infty} \delta_2 h_k(n) = -\delta_1 h_k(k-1) \leqslant 0. \tag{3.6}$$

On the other hand, for $n > k-1$, since X can make only unit downward steps, we have

$$f_k(X_n(t)+2) - 2f_k(X_n(t)+1) + f_k(X_n(t)) = 0$$

for $t < S^n_{k-1}$, and hence the inequality in (3.2) becomes the equality

$$\delta_2 h_k(n) = \mathrm{E}\exp(-2S^n_{k-1})\delta_2 h_k(k-1). \tag{3.7}$$

This in turn gives

$$\sum_{n=k-1}^{\infty} \delta_2 h_k(n) = \delta_2 h_k(k-1) \sum_{n=k-1}^{\infty} \mathrm{E}\exp(-2S^n_{k-1}), \tag{3.8}$$

which, with (3.6), implies (3.4).

To prove (3.5), observe that, if $k > 1$, then it follows from (3.6) that

$$\begin{aligned}\sum_{n=k-2}^{\infty} \delta_2 h_k(n) = h_k(k-1) - h_k(k-2) &= \int_0^{\infty} \mathrm{e}^{-t}\,\mathrm{E}[f_k(X_{k-2}(t)) - f_k(X_{k-2}(t)+1)]\,\mathrm{d}t \\ &= \mathrm{E}\int_0^{S^{k-2}_{k-1}} \mathrm{e}^{-t}[f_k(X_{k-2}(t)) - f_k(X_{k-2}(t)+1)]\,\mathrm{d}t \\ &\quad + \mathrm{E}\int_{S^{k-2}_{k-1}}^{\infty} \mathrm{e}^{-t}[f_k(X_{k-2}(t)) - f_k(X_{k-2}(t)+1)]\,\mathrm{d}t;\end{aligned}$$

but, from (1.14), $f_k(X_0(t)) - f_k(X_0(t)+1) = 0$ for $t < S^{k-2}_{k-1}$, giving

$$\begin{aligned}\sum_{n=k-2}^{\infty} \delta_2 h_k(n) &= \mathrm{E}\,\mathrm{e}^{-S^{k-2}_{k-1}}\,\mathrm{E}\int_0^{\infty} \mathrm{e}^{-s}[f_k(X_{k-1}(s)) - f_k(X_{k-1}(s)+1)]\,\mathrm{d}t \\ &> \mathrm{E}\int_0^{\infty} \mathrm{e}^{-s}[f_k(X_{k-1}(s)) - f_k(X_{k-1}(s)+1)]\,\mathrm{d}s \\ &= \sum_{n=k-1}^{\infty} \delta_2 h_k(n),\end{aligned}$$

so that (3.5) is proved.

Now, by (3.2)–(3.5), if $k > 1$, then $\delta_2 h_k(k-1) \leqslant 0$, $\delta_2 h_k(k-2) > 0$ and $\delta_2 h_k(k-1) \leqslant \delta_2 h_k(n) \leqslant \delta_2 h_k(k-2)$; if $k = 1$, then $\delta_2 h_k(0) \leqslant \delta_2 h_k(n) \leqslant 0$. Thus it suffices to show that

$$\delta_2 h_k(k-1) \geqslant -\frac{1}{\lambda_1 + 1} \qquad \text{for } k \geqslant 1 \tag{3.9}$$

and

$$\delta_2 h_k(k-2) \leqslant \frac{1}{\lambda_1 + 1} \qquad \text{for } k > 1. \tag{3.10}$$

Taking (3.9), let $V_i = \inf\{t\colon X_i(t) \neq i\}$); then, by (3.1) and using conditioning,

$$\begin{aligned}
\delta_2 h_k(k-1) = {} & \mathrm{E}\int_0^{V_{k-1}} \mathrm{e}^{-2t}(-1)\,\mathrm{d}t \\
& + \mathrm{E}\int_{V_{k-1}}^{\infty} \mathrm{e}^{-2t}[f_k(X_{k-1}(t)+2) - 2f_k(X_{k-1}(t)+1) + f_k(X_{k-1}(t))]\,\mathrm{d}t \\
= {} & -\frac{1}{\lambda_1 + k + 1} + \mathrm{E}\exp(-2V_{k-1}) \cdot \mathrm{E}\delta_2 h_k(X_{k-1}(V_{k-1})) \\
= {} & -\frac{1}{\lambda_1 + k + 1} + \mathrm{E}\exp(-2V_{k-1}) \\
& \times \left[\sum_{i=1}^{\infty} \frac{\nu_i}{\lambda_1 + k - 1}\delta_2 h_k(k-1+i) + \frac{k-1}{\lambda_1 + k - 1}\delta_2 h_k(k-2)\right] \\
\geqslant {} & -\frac{1}{\lambda_1 + k + 1} + \mathrm{E}\exp(-2V_{k-1})\sum_{i=1}^{\infty}\frac{\nu_i}{\lambda_1 + k - 1}\delta_2 h_k(k-1+i),
\end{aligned}$$

from (3.5). But now, again since X makes only unit downward jumps, we have $S_{k-1}^{k-1+i} \geqslant S_{k-1}^{k}$, almost surely, and

$$\delta_2 h_k(k-1+i) = \mathrm{E}\exp(-2S_{k-1}^{k-1+i})\delta_2 h_k(k-1) \geqslant \mathrm{E}\exp(-2S_{k-1}^{k})\delta_2 h_k(k-1) = \delta_2 h_k(k),$$

remembering that $\delta_2 h_k(k-1) \leqslant 0$. Thus, from (3.1), it follows that

$$\begin{aligned}
\delta_2 h_k(k-1) & \geqslant -\frac{1}{\lambda_1 + k + 1} + \mathrm{E}\exp(-2V_{k-1})\sum_{i=1}^{\infty}\frac{\nu_i}{\lambda_1 + k - 1}\delta_2 h_k(k) \\
& = -\frac{1}{\lambda_1 + k + 1} + \frac{\lambda_1 + k - 1}{\lambda_1 + k + 1}\frac{\lambda_1}{\lambda_1 + k - 1} \cdot \delta_2 h_k(k) \\
& = -\frac{1}{\lambda_1 + k + 1} + \frac{\lambda_1}{\lambda_1 + k + 1}\delta_2 h_k(k) \\
& = -\frac{1}{\lambda_1 + k + 1} + \frac{\lambda_1}{\lambda_1 + k + 1}\mathrm{E}\exp(-2S_{k-1}^{k})\delta_2 h_k(k-1).
\end{aligned} \tag{3.11}$$

Inequality (3.9) is now rapidly proved, once we have shown that

$$\lambda_1 e_i \leqslant i \qquad \text{for all } i \in \mathbb{N}, \tag{3.12}$$

where $e_i := \mathrm{E}\exp(-2S_{i-1}^{i})$, $i \geqslant 1$. To do so, by the Markov property and because X makes only unit downward jumps, and since $V_i \sim \exp(\lambda_1 + i)$,

$$\begin{aligned}
e_i &= \mathrm{E}\exp(-2V_i)\cdot \mathrm{E}\exp[-2(S^i_{i-1}-V_i)] \\
&= \frac{\lambda_1+i}{\lambda_1+i+2}\left\{\sum_{j=1}^{\infty}\mathrm{E}[\exp(-2(S^i_{i-1}-V_i))|X_i(V_i)=i+j]\cdot\frac{\nu_j}{\lambda_1+i}+\frac{i}{\lambda_1+i}\right\} \\
&= \frac{\lambda_1+i}{\lambda_1+i+2}\left[\sum_{j=1}^{\infty}\mathrm{E}\exp(-2S^{i+j}_{i-1})\cdot\frac{\nu_j}{\lambda_1+i}+\frac{i}{\lambda_1+i}\right] \\
&\leqslant \frac{1}{\lambda_1+i+2}[\lambda_1\mathrm{E}\exp(-2S^{i+1}_{i-1})+i] \\
&= \frac{1}{\lambda_1+i+2}[\lambda_1\mathrm{E}\exp(-2S^{i+1}_{i})\cdot\mathrm{E}\exp(-2S^{i}_{i-1})+i] \\
&= \frac{1}{\lambda_1+i+2}[\lambda_1 e_{i+1}e_i+i].
\end{aligned}$$

Hence

$$(\lambda_1+i+2)e_i \leqslant \lambda_1 e_{i+1}e_i+i,$$

which in turn implies that

$$\lambda_1 e_i - i \leqslant \lambda_1 e_{i+1}e_i-(i+2)e_i \leqslant (\lambda_1 e_{i+1}-(i+1))e_i. \tag{3.13}$$

For $i>\lambda_1$, we clearly have $\lambda_1 e_i < i$, since $e_i<1$. For $i\leqslant\lambda_1$, writing $l=[\lambda_1]+1$, (3.13) implies that

$$\lambda_1 e_i - i \leqslant (\lambda_1 e_l - l)\prod_{j=i}^{l-1} e_j < 0,$$

and so (3.12) holds for all i. Substituting this into (3.11), we have

$$\delta_2 h_k(k-1) \geqslant -\frac{1}{\lambda_1+k+1}+\frac{\lambda_1}{\lambda_1+k+1}\cdot\frac{k}{\lambda_1}\cdot\delta_2 h_k(k-1),$$

which in turn implies (3.9).

On the other hand, if $k>1$, since $\delta_2 h_k(k-2)\geqslant 0$, $\delta_2 h_k(k-1)\leqslant 0$ and $\delta_2 h_k(k-3)\leqslant \delta_2 h_k(k-2)$, it follows by the Markov property and from (3.7) that

$$\begin{aligned}
\delta_2 h_k(k-2) &= \mathrm{E}\int_0^{V_{k-2}} \mathrm{e}^{-2t}\,\mathrm{d}t + \mathrm{E}\int_{V_{k-2}}^{\infty} \mathrm{e}^{-2t}[f_k(X_{k-2}(t)+2) - 2f_k(X_{k-2}(t)+1) + f_k(X_{k-2}(t))]\,\mathrm{d}t \\
&= \frac{1}{\lambda_1 + k - 2 + 2} + \sum_{i=1}^{\infty} \frac{\nu_i}{\lambda_1 + k - 2}\delta_2 h_k(k-2+i)\cdot \mathrm{E}\,\mathrm{e}^{-2V_{k-2}} \\
&\quad + \frac{k-2}{\lambda_1 + k - 2}\cdot \delta_2 h_k(k-3)\cdot \mathrm{E}\,\mathrm{e}^{-2V_{k-2}} \\
&\leqslant \frac{1}{\lambda_1 + k} + \frac{k-2}{\lambda_1 + k - 2}\delta_2 h_k(k-2)\cdot \mathrm{E}\,\mathrm{e}^{-2V_{k-2}} \\
&= \frac{1}{\lambda_1 + k} + \frac{k-2}{\lambda_1 + k}\delta_2 h_k(k-2),
\end{aligned}$$

and (3.10) follows.

4. Applications

In this section, we show how to obtain more accurate compound Poisson approximation bounds from our estimates. As a simple illustration of what is to be gained, we consider the compound Poisson approximation to the number of k-runs of 1s in a series of independent identically distributed Bernoulli random variables ξ_i, $1 \leqslant i \leqslant n$, with $\mathrm{P}(\xi_i = 1) = p$. To avoid edge effects we treat $i + nj$ as i for $1 \leqslant i \leqslant n$, $j \in \mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}$. Define $I_i = \prod_{j=i}^{i+k-1} \xi_j$ and $W = \sum_{i=1}^{n} I_i$; then $\mathrm{E}I_i = p^k$ and $\mathrm{E}W = np^k$.

In the study of the accuracy of compound Poisson approximation to the distribution of W, Arratia *et al.* (1990) give a bound of order $nkp^{2k}(1-p)$ on the approximation error, expressed in terms of total variation distance. Under the condition $p < \frac{1}{3}$, so that the bound (1.5) can be applied, Roos (1993) improves the bound to order $kp^k \log(np^k)$. In terms of Kolmogorov distance, Theorem 4.3 of Barbour and Utev (1998) can be employed to give a bound of order $kp^k + \exp(-cnp^k)$ for some constant c (see also Eichelsbacher and Roos 1999). Here, with our new bounds on the Stein constants, we can significantly improve the error bound for Kolmogorov distance.

We use the notation of Roos (1994). Let

$$U_i = \sum_{j=i-(k-1)}^{i-1} I_j + \sum_{j=i+1}^{i+k-1} I_j,$$

the sum of I_js which strongly influence I_i, and

$$X_i = \sum_{j=i-2(k-1)}^{i-(k-1)-1} I_j + \sum_{j=i+k}^{i+2(k-1)} I_j,$$

the sum of I_js which weakly influence I_i. Then

$$\mathrm{E}U_i = \mathrm{E}X_i = 2(k-1)p^k, \qquad \mathrm{E}I_iX_i = 2(k-1)p^{2k}.$$

The parameters of the approximating compound Poisson distribution are chosen as

$$\lambda_i = \begin{cases} np^k p^{i-1}(1-p)^2, & \text{for } i = 1, 2, \ldots, k-1, \\ \dfrac{np^k p^{i-1}}{i}[2(1-p) + (2k-i-2)(1-p)^2], & \text{for } i = k, \ldots, 2k-2, \\ \dfrac{np^k p^{2k-2}}{2k-1}, & \text{for } i = 2k-1, \end{cases}$$

(see Eichelsbacher and Roos 1999) and $\lambda_1 \geqslant 2\lambda_2 \geqslant 3\lambda_3 \geqslant \ldots$ if $p \leqslant \frac{1}{3}$ or if $k \geqslant 4$ and $p \leqslant \frac{1}{2}$. Noting that I_i and U_i are independent of I_j for $j \leqslant i - 2(k-1) - 1$ or $j \geqslant i + 2(k-1) + 1$, Theorem 2 of Roos (1994), together with our improved bounds in Proposition 1.1, gives

$$d_{\mathscr{F}_{\mathrm{K}}}(\mathscr{L}(W), \mathrm{CP}(\boldsymbol{\lambda})) \leqslant \frac{1}{\lambda_1 + 1} n(6k-5)p^{2k} < \begin{cases} p, & \text{for } k = 1, \\ (6k-5)p^k(1-p)^{-2}, & \text{for } k \geqslant 2. \end{cases}$$

This simple and explicit bound, albeit for Kolmogorov rather than total variation distance, is to be compared with the previous bounds, which either grow with n or are not of optimal order unless np^k is large enough, and at best contain unspecified, and often very large, constants. Many other applications of compound Poisson approximation are given in Eichelsbacher and Roos (1999); these can be improved for Kolmogorov distance by using Proposition 1.1 in a similar way.

Acknowledgements

This work was supported by an Australian Research Council Small Grant from the University of New South Wales and by Swiss Nationalfonds Projekt no. 20-50686.97.

References

Arratia, R., Goldstein, L. and Gordon, L. (1990) Poisson approximation and the Chen–Stein method. *Stat. Sci.*, **5**, 403–434.

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

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

Barbour, A.D. and Utev, S. (1998) Solving the Stein equation in compound Poisson approximation. *Adv. Appl. Probab.*, **30**, 449–475.

Eichelsbacher, P. and Roos, M. (1999) Compound Poisson approximation for dissociated random variables via Stein's method. *Combin. Probab. Comput.*, **8**, 335–346.

Roos, M. (1993) *Stein–Chen Method for Compound Poisson Approximation*. Doctoral thesis, University of Zurich.

Roos, M. (1994) Stein's method for compound Poisson approximation: the local approach. *Ann. Appl. Probab.*, **4**, 1177–1187.

Received November 1998