Bernoulli **11**(1), 2005, 47–65

Sharp estimates in signed Poisson approximation of Poisson mixtures

JOSÉ ANTONIO ADELL* and ALBERTO LEKUONA**

*Departamento de Métodos Estadísticos, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain. E-mail: *adell@unizar.es; **lekuona@unizar.es*

We give sharp estimates in total variation and certain kinds of stop-loss metrics in signed Poisson approximation of Poisson mixtures. We provide closed-form solutions to the problem of best choice of the Poisson parameter in simple Poisson approximation with respect to the total variation distance. The important special case of the negative binomial distribution is also discussed. To obtain our results, we apply a differential calculus based on different Taylor formulae for the Poisson process which allows us to give simple unified proofs.

Keywords: Charlier polynomials; finite signed measure; Poisson approximation; Poisson mixture; probability metrics

1. Introduction

Mixtures of distributions, particularly Poisson mixtures, play an important role in many areas of applied probability and statistics, such as biology, physics, reliability and insurance. A variety of examples can be found in Johnson *et al.* (1992), Grandell (1997) and Denuit and Van Bellegem (2001), among others. With a few exceptions, the probability distribution of a Poisson mixture is quite complex to work with. It is therefore useful to approximate a Poisson mixture by a Poisson distribution, provided the mixing random variable is close to a constant. Early work in this direction goes back to Vervaat (1969) and Romanowska (1977). A broader and deeper approach may be found in Barbour (1987) and Pfeifer (1987). Recently, Roos (2003a; 2003b) has obtained sharp approximation results both in the univariate and the multivariate cases.

Several techniques have been developed in connection with general Poisson approximation problems, such as the coupling method (Serfling 1978), the semigroup technique (Deheuvels and Pfeifer 1986) and the Stein–Chen method (Chen 1975; Barbour *et al.* 1992). The degree of approximation is usually measured in terms of the total variation, the Kolmogorov or the Wasserstein distance. However, in risk theory or insurance certain stop-loss distances are used (see Rachev 1991; Roos 1999; 2001; Denuit and Van Bellegem 2001). On the other hand, a great deal of attention has been devoted to the approximation of integer-valued random variables by signed compound Poisson measures, because the degree of accuracy is better than with simple Poisson approximation (see Čekanavičius and Kruopis 2000; Barbour and Čekanavičius 2002; Roos 2002; 2003b). Such measures, coming from a Poisson-type expansion in the exponent, were apparently first considered by Kornya (1983) and Presman (1983).

The purpose of this paper is to obtain sharp estimates in signed Poisson approximation of Poisson mixtures. Although for the sake of brevity we only consider the total variation and certain kinds of stop-loss metrics, a similar treatment holds for other distances. The main feature is that we apply a differential calculus for linear operators preserving convexity which allows us to give simple unified proofs; see Adell and Lekuona (2000) for more details. It may be of interest to see this method in the case of subordinators.

Recall that a centred subordinator $X := (X(t),\ t \geqslant 0)$ is a process starting at the origin, having independent stationary increments and right-continuous non-decreasing paths, and such that $\mathrm{E}X(t) = t$, $t \geqslant 0$. The Laplace transform of $X(t)$ is given by $\mathrm{E}\mathrm{e}^{-\lambda X(t)} = \exp(-t\lambda \mathrm{E}\mathrm{e}^{-\lambda UW})$, $\lambda, t \geqslant 0$, where U is uniformly distributed on [0, 1] and W is a non-negative random variable independent of U which determines X and is called the characteristic random variable of X. Under appropriate integrability assumptions on X and smoothness assumptions on ϕ, it is established in Adell and Lekuona (2000, Corollary 1 and Proposition 4) that

$$\begin{aligned}\mathrm{E}\phi(X(t)) = {} & \sum_{k=0}^{m-1} \frac{(t-s)^k}{k!}\, \mathrm{E}\phi^{(k)}\left(X(s) + \sum_{i=1}^{k} U_i W_i \right) \\ & + \int_s^t \mathrm{E}\phi^{(m)}\left(X(u) + \sum_{i=1}^{m} U_i W_i \right) \frac{(t-u)^{m-1}}{(m-1)!}\, \mathrm{d}u, \qquad 0 \leqslant s \leqslant t, \quad m = 1, 2, \ldots,\end{aligned} \tag{1.1}$$

where $(U_i)_{i \geqslant 1}$ is a sequence of independent random variables distributed uniformly on the interval [0, 1], $(W_i)_{i \geqslant 1}$ is a sequence of independent copies of W, and $(U_i)_{i \geqslant 1}$, $(W_i)_{i \geqslant 1}$ and X are mutually independent. A striking feature is that, in many cases, a Taylor formula similar to (1.1) is valid for non-smooth functions ϕ. Therefore, in such cases, we can give a unified treatment to the problem of estimating the distance between $X(t)$ and $X(s)$ in different metrics by considering suitable sets of test functions ϕ, even if the parameter t is replaced by a random variable T. Signed Poisson approximation of Poisson mixtures fits into this framework.

One of the referees has drawn our attention to a recent paper by Roos (2003a) dealing with the same subject, in which sharp results are obtained by using similar techniques to those outlined above. A detailed comparative discussion between the methods and the results given both in Roos (2003a) and in this work will be given in the following sections. From our point of view, both papers complement each other.

The contents of this paper are organized as follows. In Section 2 we introduce the main tools, based on finite Taylor formulae for the standard Poisson process concerning arbitrary exponentially bounded functions. In Sections 3 and 4 we give sharp estimates in total variation and certain kinds of stop-loss metrics, respectively, in approximating Poisson mixtures by Charlier-type finite signed measures. In Section 5 we obtain closed-form solutions to the so-called best Poisson approximation problem, which consists of finding the Poisson distribution closest to a Poisson mixture with respect to a given distance. This problem, first posed by Serfling (1978), was developed by Deheuvels and Pfeifer (1986), Pfeifer (1987) and Deheuvels *et al.* (1989) in a more general setting, mainly from an asymptotic point of view. In the aforementioned sections, our results are stated under the weakest possible moment assumptions on the mixing random variable (see the comments after Theorem 3.1). In contrast, in Section 6 we take advantage of the integrability and limiting properties of the gamma distributions to give sharp estimates referring to the negative binomial distribution. Such estimates are uniform in the mean s of the Poisson distribution.

Throughout this paper, we shall use the following notation. A mixing random variable T is a non-negative random variable independent of the standard Poisson process $(N(s), s \geqslant 0)$. The symbol $N(T)$ stands for the corresponding Poisson mixture, the law of which is denoted by $\mathcal{P}(T)$. We denote by $\mathbb{N}$ the set of non-negative integers, and $\mathbb{N}^* := \mathbb{N}\backslash\{0\}$. For any real numbers x and y, we write $x \wedge y := \min(x, y)$, $x \vee y := \max(x, y)$, $x_+ := x \vee 0$, $x_- := (-x)_+$ and $\lfloor x \rfloor$ for the integer part of x. Every function ϕ is a real measurable function defined on $[0, \infty)$ and $\|\phi\|$ is its usual supremum-norm. We denote by 1_A the indicator function of the set A. If f and g are two positive functions, we write $f(x) \sim g(x)$ whenever $g^{-1}(x)f(x) \to 1$ as $x \to \infty$. On the other hand, for any $m \in \mathbb{N}^*$, β_m denotes a random variable having the beta density $\rho_m(\theta) := m(1-\theta)^{m-1}$, $\theta \in [0, 1]$, while $\beta_0 := 1$. Given a sequence $(U_i)_{i\geqslant 1}$ of independent and identically distributed random variables uniformly distributed on $[0, 1]$, we set $S_0 := 0$ and $S_k := U_1 + \ldots + U_k$, $k \in \mathbb{N}^*$. Finally, all of the random variables appearing under the same expectation sign are supposed to be mutually independent.

2. Taylor formulae for the Poisson process

In this section, we introduce the main tools of the paper, the most important of which is Corollary 2.1. We give two Taylor formulae for the standard Poisson process for arbitrary exponentially bounded functions, written either in terms of the forward differences of the function under consideration or in terms of the Charlier polynomials. Once such finite Taylor expansions are established, it is clear how to choose the finite signed measures to approximate Poisson mixtures.

We shall need the following ingredients. For any $\alpha \geqslant 0$, we denote by $\mathcal{E}(\alpha)$ the set of all functions ϕ such that $|\phi(x)| \leqslant C\mathrm{e}^{\alpha x}$, $x \geqslant 0$, for some constant $C \geqslant 0$. We denote by $\Delta_m\phi$ the mth forward difference of ϕ, that is,

$$\Delta_m\phi(x) := (A - I)^m\phi(x) = \sum_{k=0}^{m}(-1)^{m-k}\binom{m}{k}\phi(x+k), \qquad x \geqslant 0,\ m \in \mathbb{N}, \tag{2.1}$$

where $A\phi(x) := \phi(x+1)$, $x \geqslant 0$, I is the identity operator and $(A - I)^0 = I$. We denote by $C_m(s; x)$ the mth Charlier polynomial with respect to $N(s)$, that is,

$$C_m(s; x) := \sum_{k=0}^{m}\binom{m}{k}\binom{x}{k}k!(-s)^{-k}, \qquad x \geqslant 0,\ s > 0,\ m \in \mathbb{N}. \tag{2.2}$$

Such polynomials satisfy the orthogonality property (cf. Chihara 1978, p. 4)

$$\mathrm{E}C_k(s;\ N(s))C_m(s;\ N(s)) = \frac{m!}{s^m}\delta_{k,m}, \qquad k,\ m \in \mathbb{N}. \tag{2.3}$$

The following result relates Charlier polynomials to mth forward differences.

Lemma 2.1. *Let $\phi \in \mathcal{E}(\alpha)$ and $m \in \mathbb{N}$. Let T be a mixing random variable taking values in $(0, \infty)$ such that $\mathrm{E}\exp((\mathrm{e}^{\alpha} - 1)T) < \infty$, and let X be a non-negative random variable such that $\mathrm{E}\exp(\alpha X) < \infty$. Then, for any $k = 0, \ldots, m$, we have*

$$\mathrm{E}\phi(N(T) + X)C_m(T;\ N(T)) = (-1)^k \mathrm{E}\Delta_k\phi(N(T) + X)C_{m-k}(T;\ N(T)). \tag{2.4}$$

Proof. Let $k = 0, \ldots, m$. The formula

$$\mathrm{E}\phi(N(s))C_m(s;\ N(s)) = (-1)^k \mathrm{E}\Delta_k\phi(N(s))C_{m-k}(s;\ N(s)), \qquad s > 0, \tag{2.5}$$

may be found in Barbour *et al.* (1992, Lemma 9.1.4) or in Roos (1999, formula (6)). Applying (2.5) to the function $\tilde{\phi} \in \mathcal{E}(\alpha)$ defined by $\tilde{\phi}(x) := \mathrm{E}\phi(x + X)$, $x \geqslant 0$, and taking into account Fubini's theorem and the fact that the operators E and Δ_k commute, we obtain (2.4) with T replaced by s. By (2.1), $\Delta_m\phi \in \mathcal{E}(\alpha)$. It therefore suffices to randomize s by T in order to complete the proof of Lemma 2.1. □

We are in a position to state the following.

Theorem 2.1. *Let $t \geqslant 0$, $s > 0$ and $m \in \mathbb{N}^*$. Denote by $\gamma_m := \gamma_m(s, t) = s + (t - s)\beta_m$. Then, for any $\phi \in \mathcal{E}(\alpha)$, we have*

$$\begin{aligned}
\mathrm{E}\phi(N(t)) &= \sum_{k=0}^{m-1} \frac{(t-s)^k}{k!}\mathrm{E}\Delta_k\phi(N(s)) + \frac{(t-s)^m}{m!}\mathrm{E}\Delta_m\phi(N(\gamma_m)) \\
&= \sum_{k=0}^{m-1} \frac{(-1)^k(t-s)^k}{k!}\mathrm{E}\phi(N(s))C_k(s;\ N(s)) \\
&\quad + \frac{(-1)^m(t-s)^m}{m!}\mathrm{E}\phi(N(\gamma_m))C_m(\gamma_m;\ N(\gamma_m)).
\end{aligned}$$

Proof. Let $m \in \mathbb{N}^*$. Denote by $\mathcal{E}_m(\alpha)$ the set of m times differentiable functions ϕ such that $\phi^{(k)} \in \mathcal{E}(\alpha)$, $k = 0, \ldots, m$. For any $\phi \in \mathcal{E}_m(\alpha)$, we claim that

$$\mathrm{E}\phi(N(t)) = \sum_{k=0}^{m-1} \frac{(t-s)^k}{k!}\mathrm{E}\phi^{(k)}(N(s) + S_k) + \frac{(t-s)^m}{m!}\mathrm{E}\phi^{(m)}(N(\gamma_m) + S_m). \tag{2.6}$$

Indeed, (2.6) follows from (1.1) by observing the following. First, the standard Poisson process is a centred subordinator whose characteristic random variable W is degenerate at the point 1. Second, if $\phi \in \mathcal{E}_m(\alpha)$, the integrability conditions required in Adell and Lekuona

(2000, Corollary 1 and Proposition 4) are fulfilled. Finally, the remainder term in (1.1) can be rewritten as in (2.6) with the help of the random variable γ_m.

On the other hand, using induction, it can be checked that for any $\phi \in \mathcal{E}_m(\alpha)$ we have

$$\Delta_k \phi(x) = \mathrm{E}\phi^{(k)}(x + S_k), \qquad x \geqslant 0,\ k = 0, \ldots, m. \tag{2.7}$$

This, together with (2.6), shows that the first equality in Theorem 2.1 holds for any $\phi \in \mathcal{E}_m(\alpha)$. Let $\phi \in \mathcal{E}(\alpha)$. Since $(\phi(k),\ k \in \mathbb{N})$ are the only relevant values of ϕ, we can assume without loss of generality that ϕ is continuous. Therefore (cf. Adell and Sangüesa 2004) there is a sequence of functions $(\phi_n,\ n \in \mathbb{N}) \subseteq \mathcal{E}_m(\alpha)$ such that

$$\lim_{n\to\infty} \Delta_k \phi_n(x) = \Delta_k \phi(x), \qquad x \geqslant 0,\ k = 0, \ldots, m.$$

Hence, the first equality in Theorem 2.1 follows from dominated convergence, while the second readily follows from Lemma 2.1. The proof is complete. $\square$

Setting $\phi = 1_{\{n\}}$, $n \in \mathbb{N}$, in the second equality in Theorem 2.1, we obtain

$$\begin{aligned} P(N(t) = n) &= \sum_{k=0}^{m-1} \frac{(-1)^k (t-s)^k}{k!} C_k(s;\, n) P(N(s) = n) \\ &\quad + \frac{(-1)^m (t-s)^m}{m!} \mathrm{E}1_{\{n\}}(N(s + (t-s)\beta_m)) C_m(s + (t-s)\beta_m;\, N(s + (t-s)\beta_m)). \end{aligned} \tag{2.8}$$

We emphasize that (2.8) coincides, up to changes of notation, with the formula given by Roos (2003a, Lemma 2 and formula (9)) which, according to the author, constitutes the main tool in the argument of his paper. In turn, the formulae in Theorem 2.1 can be derived from (2.8) by integration. Therefore, as far as signed Poisson approximation of Poisson mixtures is concerned, the main tools developed both in Roos (2003a) and in this paper are, up to changes of notation, essentially equivalent. Finally, a multivariate version of (2.8) can be found in Roos (2003b, Lemma 1).

Let $s > 0$ be fixed and let $(T_n,\ n \in \mathbb{N})$ be a sequence of mixing random variables converging to s as $n \to \infty$. For any $k, n \in \mathbb{N}$, we write $\mu_k(n) := \mu_{k,s}(n) = \mathrm{E}(T_n - s)^k$, whenever the expectation exists. Let $m, n \in \mathbb{N}$ and assume that $\mathrm{E}|T_n - s|^m < \infty$. In view of Theorem 2.1 or (2.8), we consider the Charlier-type finite signed measure $\nu_m^{(n)}$ on $\mathbb{N}$ defined by

$$\nu_m^{(n)}(\{l\}) := \nu_{m,s}^{(n)}(\{l\}) = \frac{\mathrm{e}^{-s} s^l}{l!} \sum_{k=0}^{m} \frac{(-1)^k \mu_k(n)}{k!} C_k(s;\, l), \qquad l \in \mathbb{N}. \tag{2.9}$$

Observe that $\nu_0^{(n)} = \mathcal{P}(s)$. Also, $\nu_1^{(n)} = \mathcal{P}(s)$, provided that $\mathrm{E}T_n = s$. The following result, which is an immediate consequence of Theorem 2.1, will be crucial throughout the rest of this paper.

Corollary 2.1. *Let $s > 0$ and $m, n \in \mathbb{N}$. Assume that $\mathrm{E}|T_n - s|^m < \infty$. Then, for any $\phi \in \mathcal{E}(\alpha)$ such that $\|\Delta_m \phi\| < \infty$, we have*

$$\mathrm{E}\phi(N(T_n)) - \int_{\mathbb{N}} \phi(l)\nu_m^{(n)}(\mathrm{d}l) = \frac{1}{m!}\mathrm{E}(T_n - s)^m(\Delta_m\phi(N(s + (T_n - s)\beta_m)) - \Delta_m\phi(N(s))).$$

In the following sections, we shall need to bound the term on the right-hand side in Corollary 2.1. The following auxiliary result goes in this direction.

Lemma 2.2. *Let* $s, t \geqslant 0$. *If* $\|\phi\| < \infty$, *then*

$$|\mathrm{E}\phi(N(t)) - \mathrm{E}\phi(N(s))| \leqslant (\|\phi_+\| + \|\phi_-\|)(1 - \mathrm{e}^{-|t-s|}).$$

Proof. Denote by $w(\phi;\, \cdot)$ the usual first modulus of continuity of ϕ, that is,

$$w(\phi;\, \epsilon) := \sup\{|\phi(x) - \phi(y)|\colon x,\, y \geqslant 0,\, |x - y| \leqslant \epsilon\}, \qquad \epsilon \geqslant 0.$$

Using the inequalities $|\phi(x) - \phi(y)| \leqslant w(\phi;\, |x - y|) \leqslant \|\phi_+\| + \|\phi_-\|$, $x, y \geqslant 0$, we have

$$\begin{aligned}|\mathrm{E}\phi(N(t)) - \mathrm{E}\phi(N(s))| &\leqslant \mathrm{E}w(\phi;\, |N(t) - N(s)|) = Ew(\phi;\, N(|t - s|)) \\ &\leqslant (\|\phi_+\| + \|\phi_-\|)P(N(|t - s|) \geqslant 1),\end{aligned}$$

thus completing the proof. □

3. Total variation distance

If μ and ν are finite signed measures on $\mathbb{N}$, the total variation distance between μ and ν is defined by

$$d_{\mathrm{TV}}(\mu, \nu) := \sup_{A\subseteq\mathbb{N}}\left|\int_{\mathbb{N}} 1_A(l)(\mu - \nu)(\mathrm{d}l)\right|.$$

In this section, we give upper bounds and exact estimates – up to a remainder term – for $\mathrm{d}_{\mathrm{TV}}(\mathcal{P}(T_n), \nu_m^{(n)})$. As an application, we estimate the total variation distance between two Poisson distributions with different means.

Theorem 3.1. *Let* $s > 0$ *and* $m, n \in \mathbb{N}$.

(a) *If* $\mathrm{E}|T_n - s|^m < \infty$, *then*

$$d_{\mathrm{TV}}(\mathcal{P}(T_n), \nu_m^{(n)}) \leqslant \frac{2^m}{m!}\,\mathrm{E}|T_n - s|^m(1 - \mathrm{e}^{-|T_n - s|\beta_m}).$$

(b) *If* $\mathrm{E}|T_n - s|^{m+1} < \infty$, *then*

$$\left|d_{\mathrm{TV}}(\mathcal{P}(T_n), \nu_m^{(n)}) - \frac{|\mu_{m+1}(n)|}{(m+1)!\,2}\,\mathrm{E}|C_{m+1}(s;\, N(s))|\right| \leqslant \frac{2^{m+1}}{(m+1)!}\,\mathrm{E}|T_n - s|^{m+1}(1 - \mathrm{e}^{-|T_n - s|\beta_{m+1}}).$$

Proof. (a) For any $m \in \mathbb{N}$ and $A \subseteq \mathbb{N}$, let $b_m(A) := \|(\Delta_m 1_A)_+\| + \|(\Delta_m 1_A)_-\|$. From Corollary 2.1, Lemma 2.2 and Fubini's theorem, we have

$$\begin{aligned} d_{\mathrm{TV}}(\mathcal{P}(T_n), \nu_m^{(n)}) &\leqslant \frac{1}{m!} \sup_{A \subseteq \mathbb{N}} |\mathrm{E}(T_n - s)^m (\Delta_m 1_A(N(s + (T_n - s)\beta_m)) - \Delta_m 1_A(N(s)))| \\ &\leqslant \frac{1}{m!} \mathrm{E}|T_n - s|^m (1 - \mathrm{e}^{-|T_n - s|\beta_m}) \sup_{A \subseteq \mathbb{N}} b_m(A). \end{aligned} \tag{3.1}$$

It is readily seen from (2.1) that $b_0(A) = 1$ and $b_m(A) \leqslant 2^{m-1} + 2^{m-1} = 2^m$, $m \in \mathbb{N}^*$, $A \subseteq \mathbb{N}$. Thus, part (a) follows from (3.1).

(b) Let $m \in \mathbb{N}$ and $A_m := \mathbb{N} \cap \{x \geqslant 0\colon C_m(s; x) \geqslant 0\}$. By (2.5) and (2.3), we see that

$$\begin{aligned} \sup_{A \subseteq \mathbb{N}} |\mathrm{E}\Delta_m 1_A(N(s))| &= \sup_{A \subseteq \mathbb{N}} |\mathrm{E}1_A(N(s)) C_m(s; N(s))| \\ &= |\mathrm{E}1_{A_m}(N(s)) C_m(s; N(s))| \\ &= \tfrac{1}{2} \mathrm{E}|C_m(s; N(s))|. \end{aligned} \tag{3.2}$$

On the other hand, let $\phi \in \mathcal{E}(\alpha)$ be such that $\|\Delta_{m+1}\phi\| < \infty$. By considering one more term in the expansion given in Corollary 2.1, we can write

$$\begin{aligned} \mathrm{E}\phi(N(T_n)) - \int_{\mathbb{N}} \phi(l) \nu_m^{(n)}(\mathrm{d}l) \;=\; & \frac{\mu_{m+1}(n)}{(m+1)!} \mathrm{E}\Delta_{m+1}\phi(N(s)) \\ & + \frac{1}{(m+1)!} \mathrm{E}(T_n - s)^{m+1} \big(\Delta_{m+1}\phi(N(s + (T_n - s)\beta_{m+1})) \\ & - \Delta_{m+1}\phi(N(s))\big). \end{aligned} \tag{3.3}$$

Applying (3.3) to $\phi = 1_A$, $A \subseteq \mathbb{N}$, and recalling (3.2), we obtain as in part (a) that

$$\begin{aligned} & \left| d_{\mathrm{TV}}(\mathcal{P}(T_n), \nu_m^{(n)}) - \frac{|\mu_{m+1}(n)|}{(m+1)!\,2} \mathrm{E}|C_{m+1}(s; N(s))| \right| \\ & \quad \leqslant \sup_{A \subseteq \mathbb{N}} \left| \mathrm{E}1_A(N(T_n)) - \int_{\mathbb{N}} 1_A(l) \nu_m^{(n)}(\mathrm{d}l) - \frac{\mu_{m+1}(n)}{(m+1)!} \mathrm{E}\Delta_{m+1} 1_A(N(s)) \right| \\ & \quad \leqslant \frac{2^{m+1}}{(m+1)!} \mathrm{E}|T_n - s|^{m+1} (1 - \mathrm{e}^{-|T_n - s|\beta_{m+1}}). \end{aligned}$$

This shows (b) and completes the proof. □

Observe that the estimates in Theorem 3.1 are given under the weakest possible moment assumptions on T_n. This follows from the fact that the finiteness of the mth moment of T_n is needed to define the signed measure $\nu_m^{(n)}$ and, at the same time, the error bound is finite without any further assumptions on T_n. However, it should be mentioned that estimates under the weakest possible moment assumptions on T_n are not always good estimates. This can be seen by considering the case $m = 0$ in Theorem 3.1(a), which gives us $\mathrm{d}_{\mathrm{TV}}(\mathcal{P}(T_n), \mathcal{P}(s)) \leqslant \mathrm{E}(1 - \exp(-|T_n - s|))$.

Similar estimates to those given in Theorem 3.1 have been shown by Roos (2003a, Theorem 2), who obtains

$$d_{\mathrm{TV}}(\mathcal{P}(T_n), \nu_m^{(n)}) \leqslant \min\left\{ \frac{U_1^{(m+1)}}{m!\,2} \mathrm{E}\left| \int_s^{T_n} \frac{|T_n - y|^m}{y^{(m+1)/2}} \,\mathrm{d}y \right|, \frac{2^m}{(m+1)!} \mathrm{E}|T_n - s|^{m+1} \right\}, \tag{3.4}$$

where $U_1^{(m+1)}$ is a positive constant. For $m = 2, 3$ and whenever $\mathrm{E}T_n = s$, Roos (2003a, formulae (23) and (45)) has shown that the first upper bound in (3.4) can in turn be bounded above by

$$\frac{U_1^{(3)}}{2\sqrt{s}} \mathrm{E}\left(\frac{|T_n - s|^3}{(\sqrt{T_n} + \sqrt{s})^2} \right) \quad \text{and} \quad \frac{U_1^{(4)}}{12s} \mathrm{E}\left(\frac{(T_n - s)^4}{T_n + 2s} \right), \tag{3.5}$$

respectively. The quantities in (3.5) are finite if $\mathrm{E}(T_n - s)^2 < \infty$ and $\mathrm{E}|T_n - s|^3 < \infty$, respectively. This means that, in the cases at hand, the estimates in (3.4) are also given under the weakest possible moment assumptions on T_n.

Using a semigroup approach, Pfeifer (1987) obtained for $m = 1$ the same main term as in Theorem 3.1(b), with different error bound. On the other hand, exact values of $\mathrm{E}|C_m(s; N(s))|$ for $m = 1, 2$, in terms of the integer parts of the roots of $C_m(s; x)$ were obtained by Deheuvels and Pfeifer (1986). Finally, sharp upper bounds for $\mathrm{E}|C_m(s; N(s))|$, $m \in \mathbb{N}$, can be found in Roos (2003a, Lemmas 3 and 4) and Roos (1999, 2001).

As an application of Theorem 3.1, we give the following.

Corollary 3.1. *For any $0 < s \leqslant t$, we have*

$$\begin{aligned} \mathrm{d}_{\mathrm{TV}}(\mathcal{P}(t), \mathcal{P}(s)) &\leqslant \min\left\{ 1 - \mathrm{e}^{-(t-s)}, \int_s^t P(N(y) = \lfloor y \rfloor) \mathrm{d}y \right\} \\ &\leqslant \min\left\{ 1 - \mathrm{e}^{-(t-s)}, (t-s) P(N(s) = \lfloor s \rfloor) \right\} \end{aligned} \tag{3.6}$$

and

$$|d_{\mathrm{TV}}(\mathcal{P}(t), \mathcal{P}(s)) - (t-s) P(N(s) = \lfloor s \rfloor)| \leqslant (t-s)^2. \tag{3.7}$$

Proof. As shown by Deheuvels and Pfeifer (1986), we have $\mathrm{E}|C_1(s; N(s))| = 2P(N(s) = \lfloor s \rfloor)$, $s > 0$. Therefore, by Theorem 2.1 and (3.2), we have

$$\sup_{A\subseteq\mathbb{N}}|\mathrm{E}1_A(N(t)) - \mathrm{E}1_A(N(s))| \leq (t-s)\int_0^1 \sup_{A\subseteq\mathbb{N}}|\mathrm{E}\Delta_1 1_A(N(s+(t-s)\theta))|\mathrm{d}\theta$$

$$= \frac{t-s}{2}\int_0^1 \mathrm{E}|C_1(s+(t-s)\theta;\ N(s+(t-s)\theta))|\mathrm{d}\theta$$

$$= \int_s^t P(N(y) = \lfloor y \rfloor)\,\mathrm{d}y \leq (t-s)P(N(s) = \lfloor s \rfloor),$$

the last inequality because the function $g(y) := P(N(y) = \lfloor y \rfloor)$, $y \geq 0$, is non-increasing. This shows the second upper bound in (3.6). The first upper bound in (3.6) follows from Theorem 3.1(a) by choosing $T_n = t$ and $m = 0$. Similarly, inequality (3.7) follows from Theorem 3.1(b). □

In the setting of Corollary 3.1, Roos (2003a, formula (5)) gives the bound

$$d_{\mathrm{TV}}(\mathcal{P}(t), \mathcal{P}(s)) \leq \min\left\{ t-s,\ \sqrt{\frac{2}{\mathrm{e}}}(\sqrt{t} - \sqrt{s}) \right\}. \tag{3.8}$$

It can be shown by calculus that $P(N(y) \leq \lfloor y \rfloor) \leq (2ey)^{-1/2}$, $y > 0$, with equality if and only if $y = \frac{1}{2}$. This implies that the first upper bound in (3.6) is better than that in (3.8). However, the more explicit estimate given by the second bound in (3.6) is not uniformly better than that in (3.8). Notwithstanding this fact, the second upper bound in (3.6) is less than 1 and asymptotically optimal in the sense that if t is close to s and s tends to infinity, then we have from (3.7) that

$$d_{\mathrm{TV}}(\mathcal{P}(t), \mathcal{P}(s)) \sim (t-s)P(N(s) = \lfloor s \rfloor) \sim (t-s)\frac{1}{\sqrt{2\pi s}}.$$

4. Stop-loss distances

If μ and ν are finite signed measures on $\mathbb{N}$ and $r \in \mathbb{N}$, we define

$$d_{r,\infty}(\mu, \nu) := \frac{1}{r!}\sup_{n\in\mathbb{N}}\left|\int_{\mathbb{N}}(l-n)_+^r(\mu-\nu)(\mathrm{d}l)\right|, \quad d_{r,1}(\mu, \nu) := \frac{1}{r!}\sum_{n=0}^{\infty}\left|\int_{\mathbb{N}}(l-n)_+^r(\mu-\nu)(\mathrm{d}l)\right|,$$

where it is understood that $(x-a)_+^0 = 1_{[a,\infty)}(x)$, $a, x \geq 0$. If μ and ν are probability measures, $d_{0,\infty}$ is the Kolmogorov distance, $d_{0,1}$ is the Wasserstein distance, while for $r \in \mathbb{N}^*$, $d_{r,\infty}$ and $d_{r,1}$ are called stop-loss distances (see Rachev (1991)). For the sake of brevity, we shall only give in this section estimates for $d_{r,1}(\mathcal{P}(T_n), \nu_m^{(n)})$ analogous to those stated in Theorem 3.1(b).

Let $s > 0$, $m \in \mathbb{N}^*$ and $r \in \mathbb{N}$. Denote by

$$c_r(m, s) := \sum_{n=0}^{\infty}|\mathrm{E}\Delta_m\phi_{r,n}(N(s))|, \qquad \phi_{r,n}(x) := \frac{1}{r!}(x-n)_+^r, \quad x \geq 0. \tag{4.1}$$

It turns out that $c_r(\cdot, \cdot)$ are the constants appearing in the main terms of the approximation. Such constants are estimated in the following auxiliary result.

Lemma 4.1. *Let $s > 0$ and $m \in \mathbb{N}^*$. We have*

(a) $c_0(m, s) = \mathrm{E}|C_{m-1}(s; N(s))|$.

(b) If $r \in \mathbb{N}^*$, then

$$c_r(m, s) = \frac{1}{(r-m)!} \mathrm{E}(N(s) + S_m)^{r-m}(\lfloor (N(s) + S_m)\beta_{r-m} \rfloor + 1), \qquad \text{if } m \leqslant r,$$

$$c_1(m, s) = c_0(m-1, s), \qquad \text{if } m > r = 1,$$

$$c_r(r+1, s) = 1,$$

and

$$c_r(m, s) \leqslant c_0(m-r, s), \qquad \text{if } m > r + 1 > 2.$$

Proof. Part (a) follows from Roos (1999, Corollary 2).

Assume that $r \in \mathbb{N}^*$. If $m \leqslant r$, we have from (2.7) that

$$\mathrm{E}\Delta_m \phi_{r,n}(N(s)) = \frac{1}{(r-m)!} \mathrm{E}(N(s) + S_m - n)_+^{r-m}. \tag{4.2}$$

On the other hand, if X is a non-negative random variable, we claim that

$$\sum_{n=0}^{\infty} \mathrm{E}(X - n)_+^r = EX^r(\lfloor X\beta_r \rfloor + 1). \tag{4.3}$$

Indeed, it can be checked that $(x-n)_+^r = x^r \mathrm{E}1_{[n,\infty)}(x\beta_r)$, $x \geqslant 0$, $n \in \mathbb{N}$, thus implying that

$$\sum_{n=0}^{\infty} (x-n)_+^r = x^r \mathrm{E}(\lfloor x\beta_r \rfloor + 1), \qquad x \geqslant 0.$$

This, together with Fubini's theorem, shows claim (4.3). Therefore, the first equality in (b) follows from (4.2) and (4.3).

If $m > r$, we again have from (2.7) that

$$\mathrm{E}\Delta_m \phi_{r,n}(N(s)) = \mathrm{E}\Delta_{m-r} 1_{[n,\infty)}(N(s) + S_r) = \mathrm{E}\Delta_{m-r} 1_{[n,\infty)}(N(s) + \lfloor S_r \rfloor). \tag{4.4}$$

Thus, the remaining three statements in part (b) follow from (4.4). □

We are now in a position to state the following.

Theorem 4.1. *Let $s > 0$ and $m, n, r \in \mathbb{N}$. Assume that $\mathrm{E}|T_n - s|^{(m \vee r)+1} < \infty$. Then,*

(a) *if $m + 1 \leqslant r$, then*

$$\left| d_{r,1}(\mathcal{P}(T_n), \nu_m^{(n)}) - \frac{|\mu_{m+1}(n)|}{(m+1)!} c_r(m+1, s) \right| \leq \sum_{k=m+2}^{r} \frac{|\mu_k(n)|}{k!} c_r(k, s) + \frac{\mathrm{E}|T_n - s|^{r+1}}{(r+1)!};$$

(b) *if* $m+1 > r$, *then*

$$\left| d_{r,1}(\mathcal{P}(T_n), \nu_m^{(n)}) - \frac{|\mu_{m+1}(n)|}{(m+1)!} c_r(m+1, s) \right| \leq \frac{2^{m+1-r}}{(m+1)!} \mathrm{E}|T_n - s|^{m+1}(1 - \mathrm{e}^{-|T_n - s|\beta_{m+1}}).$$

Proof. (a) Let $i \in \mathbb{N}$ and assume that $m+1 \leq r$. By (2.7), we have

$$\Delta_{r+1}\phi_{r,i}(x) = \mathrm{E}\Delta_1 1_{[i,\infty)}(x + S_r) = \mathrm{E}1_{[i-1,i)}(x + S_r), \qquad x \geq 0. \tag{4.5}$$

Hence, applying Corollary 2.1 to $\phi_{r,i}$, we obtain from (4.5) that

$$\begin{aligned}
&\mathrm{E}\phi_{r,i}(N(T_n)) - \int_{\mathbb{N}} \phi_{r,i}(l)\nu_m^{(n)}(\mathrm{d}l) \\
&= \frac{\mu_{m+1}(n)}{(m+1)!} \mathrm{E}\Delta_{m+1}\phi_{r,i}(N(s)) + \sum_{k=m+2}^{r} \frac{\mu_k(n)}{k!} \mathrm{E}\Delta_k\phi_{r,i}(N(s)) \\
&\quad + \frac{1}{(r+1)!}\mathrm{E}(T_n - s)^{r+1} 1_{[i-1,i)}(N(s + (T_n - s)\beta_{r+1}) + S_r).
\end{aligned} \tag{4.6}$$

Thus, part (a) follows from (4.6) and (4.1) as in the proof of Theorem 3.1(b).

(b) Let $i \in \mathbb{N}$ and assume that $m+1 > r$. Applying Corollary 2.1 to $\phi_{r,i}$ and taking into account (4.4), we have

$$\begin{aligned}
\mathrm{E}\phi_{r,i}(N(T_n)) - \int_{\mathbb{N}} \phi_{r,i}(l)\nu_m^{(n)}(\mathrm{d}l) &= \frac{\mu_{m+1}(n)}{(m+1)!}\mathrm{E}\Delta_{m+1}\phi_{r,i}(N(s)) \\
&\quad + \frac{1}{(m+1)!}\mathrm{E}(T_n - s)^{m+1} h_i(T_n, m, s),
\end{aligned} \tag{4.7}$$

where

$$h_i(T_n, m, s) := \Delta_{m+1-r}1_{[i,\infty)}(N(s + (T_n - s)\beta_{m+1}) + \lfloor S_r \rfloor) - \Delta_{m+1-r}1_{[i,\infty)}(N(s) + \lfloor S_r \rfloor).$$

Let $u, v > 0$ and $k \in \mathbb{N}^*$. Using (2.1) and Theorem 3.1(a) with $m = 0$, we have

$$\begin{aligned}
&\sum_{i=0}^{\infty} |\mathrm{E}\Delta_k 1_{[i,\infty)}(N(u)) - \mathrm{E}\Delta_k 1_{[i,\infty)}(N(v))| \\
&\quad = \sum_{i=0}^{\infty} |\mathrm{E}\Delta_{k-1} 1_{\{i-1\}}(N(u)) - \mathrm{E}\Delta_{k-1} 1_{\{i-1\}}(N(v))| \\
&\quad \leqslant \sum_{j=0}^{k-1} \binom{k-1}{j} \sum_{i=0}^{\infty} |\mathrm{E}1_{\{i-1-j\}}(N(u)) - \mathrm{E}1_{\{i-1-j\}}(N(v))| \\
&\quad \leqslant 2^{k-1} 2 d_{\mathrm{TV}}(\mathcal{P}(u), \mathcal{P}(v)) \leqslant 2^k (1 - \mathrm{e}^{-|u-v|}). \qquad (4.8)
\end{aligned}$$

Therefore, part (b) follows from (4.1), (4.7) and (4.8). □

Stop-loss distances of a different kind in a general Poisson approximation setting have been considered by Roos (1999, 2001, 2003a). Specifically, if μ and ν are finite signed measures on $\mathbb{N}$, Roos (2003a, formula (10)) defines the distances

$$d^*_{r,\infty}(\mu, \nu) := \sup_{n\in\mathbb{N}} \left| \int_{\mathbb{N}} \Delta^{-r} 1_{\{n\}}(l)(\mu - \nu)(\mathrm{d}l) \right|, \; d^*_{r,1}(\mu, \nu) := \sum_{n=0}^{\infty} \left| \int_{\mathbb{N}} \Delta^{-r} 1_{\{n\}}(l)(\mu - \nu)(\mathrm{d}l) \right|,$$

where $r \in \mathbb{N}^*$, $\Delta^{-1} f(l) := -(f(0) + \ldots + f(l))$ and $\Delta^{-(r+1)} f(l) := \Delta^{-r}(\Delta^{-1} f)(l)$, $l \in \mathbb{N}$, $r \in \mathbb{N}^*$. Using induction, it can be seen that

$$\Delta^{-r} 1_{\{n\}}(l) = (-1)^r \binom{l-n+r-1}{r-1} 1_{[n,\infty)}(l), \qquad l, n \in \mathbb{N}, \; r \in \mathbb{N}^*.$$

Thus, $d^*_{r,\infty}$ and $d^*_{r,1}$ measure the l_∞ and the l_1 distances between truncated factorial moments of μ and ν, respectively. Whenever $m + 1 \geqslant r$, upper bounds for $d^*_{r,\infty}(\mathcal{P}(T_n), \nu_m^{(n)})$ and $d^*_{r,1}(\mathcal{P}(T_n), \nu_m^{(n)})$ analogous to those given in (3.4) can be found in Roos (2003a, Theorem 2).

5. Best Poisson approximation

In this section, we give closed-form solutions to the problem of best choice of the Poissonian mean in the setting of simple Poisson approximation of Poisson mixtures. Attention will be focused on the total variation distance.

Let $s > 0$ be fixed and assume that the sequence of mixing random variables $(T_n, \; n \in \mathbb{N})$ satisfies $\mathrm{E}T_n = s$, $n \in \mathbb{N}$. For any real a, let $P_a(s; x)$ be the quadratic polynomial

$$P_a(s; x) := C_2(s; x) - \frac{a}{s} C_1(s; x) = \frac{1}{s^2}(x^2 - (2s + 1 - a)x + s(s - a)), \qquad (5.1)$$

where the last equality follows from (2.2). The roots of $P_a(s; x)$ are

$$r_i(a) := s + \frac{1-a}{2} + (-1)^i \sqrt{s + \left(\frac{1-a}{2}\right)^2}, \qquad i = 1, 2. \qquad (5.2)$$

It turns out that the integer parts of these roots are relevant to the problem at hand. It is easily verified that $r_1(a)$ is a decreasing function, $r_1(a) < 0$ for $a > s$, and

$$\lim_{a\to-\infty} r_1(a) = s, \qquad \lim_{a\to\infty} r_1(a) = -\infty. \tag{5.3}$$

Similarly, $r_2(a)$ is decreasing, $r_2(a) \geqslant s$ for any real a, and

$$\lim_{a\to-\infty} r_2(a) = \infty, \qquad \lim_{a\to\infty} r_2(a) = s. \tag{5.4}$$

Finally, for any real a, denote by $n_i(a) := \lfloor r_i(a) \rfloor$, $i = 1, 2$, and by

$$j_s(a) := \mathrm{E}(1_{\{n_1(a)\}} - 1_{\{n_2(a)\}})(N(s))\Big(C_1(s; N(s)) - \frac{a}{s}\Big). \tag{5.5}$$

Theorem 5.1. *Assume that $\mu_2(n) \to 0$ as $n \to \infty$. Then, for any $a \leqslant (2s^2)/\mu_2(n)$, we have*

$$\begin{aligned}\Big| d_{\mathrm{TV}}\Big(N(T_n), N\Big(s - \frac{a}{2s}\mu_2(n)\Big)\Big) - \frac{\mu_2(n)}{2} j_s(a)\Big| \\ \leqslant 2\mathrm{E}|T_n - s|^2(1 - \mathrm{e}^{-|T_n - s|\beta_2}) + \Big(\frac{a}{2s}\,\mu_2(n)\Big)^2.\end{aligned} \tag{5.6}$$

Moreover, $j_s(a)$ is a piecewise linear convex function which attains its minimum at

$$a_* := \inf\left\{a \leqslant s\colon \quad \frac{s^{n_1(a)}}{n_1(a)!} \leqslant \frac{s^{n_2(a)}}{n_2(a)!}\right\}. \tag{5.7}$$

Proof. Let $n \in \mathbb{N}$ and let ϕ be a function such that $\|\Delta_2\phi\| < \infty$. Using a Taylor expansion around $N(s)$ as in Corollary 2.1 and Theorem 2.1, we obtain

$$\mathrm{E}\phi(N(T_n)) - \mathrm{E}\phi\Big(N\Big(s - \frac{a}{2s}\mu_2(n)\Big)\Big) = \frac{\mu_2(n)}{2}\mathrm{E}\phi(N(s))P_a(s; N(s)) + R_n(\phi), \tag{5.8}$$

where $P_a(s; x)$ is the polynomial defined in (5.1) and

$$\begin{aligned}R_n(\phi) := {}& \frac{1}{2}\,\mathrm{E}(T_n - s)^2(\Delta_2\phi(N(s + (T_n - s)\beta_2)) - \Delta_2\phi(N(s))) \\ & + \frac{a}{2s}\,\mu_2(n)\mathrm{E}\Big(\Delta_1\phi\Big(N\Big(s - \frac{a}{2s}\,\mu_2(n)\beta_1\Big)\Big) - \Delta_1\phi(N(s))\Big).\end{aligned} \tag{5.9}$$

Recalling (5.1) and (5.2), and using (2.3) and (2.5), we have

$$\sup_{A\subseteq\mathbb{N}}|\mathrm{E}1_A(N(s))P_a(s; N(s))| = -\mathrm{E}1_{[n_1(a)+1, n_2(a)]}(N(s))P_a(s; N(s)) = j_s(a). \tag{5.10}$$

On the other hand, by Lemma 2.2, we have as in the proof of Theorem 3.1(a) that

$$\sup_{A\subseteq\mathbb{N}}|R_n(1_A)| \leqslant 2\mathrm{E}|T_n - s|^2(1-\mathrm{e}^{-|T_n-s|\beta_2}) + \frac{|a|}{s}\,\mu_2(n)\mathrm{E}(1-\mathrm{e}^{-(|a|/2s)\,\mu_2(n)\beta_1})$$

$$\leqslant 2\mathrm{E}|T_n - s|^2(1-\mathrm{e}^{-|T_n-s|\beta_2}) + \left(\frac{a}{2s}\,\mu_2(n)\right)^2. \tag{5.11}$$

As in the proof of Theorem 3.1(b), (5.6) follows from (5.8)–(5.11).

To show the last statement in the theorem, observe that by (5.5) the function $j_s(a)$ is piecewise linear with slope given by

$$m_s(a) := -\frac{1}{s}(P(N(s) = n_1(a)) - P(N(s) = n_2(a))).$$

On the other hand, it follows from (5.3) and (5.4) that $n_1(a)$ and $n_2(a)$ are decreasing functions with $n_1(a) \leqslant \lfloor s\rfloor \leqslant \lfloor s\rfloor + 1 \leqslant n_2(a)$, for any real a. Let $a \leqslant a'$. Since the function $p_s(n) := P(N(s) = n)$, $n \in \mathbb{N}$, is increasing in $\{0, \ldots, \lfloor s\rfloor\}$ and decreasing in $\{\lfloor s\rfloor, \lfloor s\rfloor + 1, \ldots\}$, we have

$$P(N(s) = n_1(a')) \leqslant P(N(s) = n_1(a)), \qquad P(N(s) = n_2(a)) \leqslant P(N(s) = n_2(a')).$$

This implies that $m_s(a) \leqslant m_s(a')$ and shows the convexity of $j_s(a)$. Finally, since $n_1(a) < 0$ for $a > s$, we see that $m_s(a) > 0$ for $a > s$. Therefore, the convex function $j_s(a)$ attains its minimum at

$$a_* = \inf\{a \leqslant s : m_s(a) \geqslant 0\} = \inf\left\{a \leqslant s\colon \frac{s^{n_1(a)}}{n_1(a)!} \leqslant \frac{s^{n_2(a)}}{n_2(a)!}\right\}.$$

The proof of the theorem is complete. □

Thanks to (5.7), we can obtain explicit values of a_* for small values of s, as was done in Deheuvels *et al.* (1989). Also, (5.7) provides an algorithm to evaluate a_* in a finite number of steps. This algorithm can be implemented in a system for mathematical computation such as MAPLE. Actually, it is not hard to see that the sequences of points $\{r_1^{-1}(i), i \leqslant \lfloor s\rfloor - \lfloor\lfloor s\rfloor s^{-1}\rfloor\}$ and $\{r_2^{-1}(i), i \geqslant \lfloor s\rfloor + 1\}$ are the interpolation points of j_s, where it is understood that i is an arbitrary integer. Since j_s is convex, we can compute a_* as $j_s(a_*) = j_s(r_1^{-1}(K)) \wedge j_s(r_2^{-1}(M))$, where

$$r_1^{-1}(K) := \min\{r_1^{-1}(i) : j_s(r_1^{-1}(i)) \leqslant j_s(r_1^{-1}(i+1)),\ i \leqslant \lfloor s\rfloor - \lfloor\lfloor s\rfloor s^{-1}\rfloor - 1\}$$

and

$$r_2^{-1}(M) := \min\{r_2^{-1}(i)\colon j_s(r_2^{-1}(i)) \leqslant j_s(r_2^{-1}(i+1)),\ i \geqslant \lfloor s\rfloor + 1\}.$$

6. The negative binomial distribution

Recall that a gamma process $(X(t),\ t \geqslant 0)$ is a centred subordinator such that $X(t)$ has the gamma density

$$\rho_t(\theta) := \frac{\theta^{t-1}\mathrm{e}^{-\theta}}{\Gamma(t)}, \qquad \theta > 0,\ t > 0.$$

For any $b \geqslant 1$ and $s > 0$, we consider the mixing random variable $T_b(s) := b^{-1}X(bs)$. Observe that $\mathrm{E}T_b(s) = s$. It is well known that the random variable $N(T_b(s))$ has the negative binomial distribution given by

$$P(N(T_b(s)) = k) = \binom{bs+k-1}{k}\left(\frac{1}{1+b}\right)^k\left(\frac{b}{1+b}\right)^{bs}, \qquad k \in \mathbb{N}.$$

For notational simplicity, we shall consider in this section a continuous parameter $b \geqslant 1$ and denote by $\nu_m^{(b)}$ the finite signed measure defined in (2.9), where $n \in \mathbb{N}$ is replaced by $b \geqslant 1$. Our aim is to obtain uniform estimates in the mean s for $d_{\mathrm{TV}}(\mathcal{P}(T_b(s)), \nu_m^{(b)})$ as $b \to \infty$. To do this, we bound the remainder term in Corollary 2.1 in a different way than done in the previous sections. It turns out that in approximating $\mathcal{P}(T_b(s))$ by $\nu_{2m}^{(b)}$ and $\nu_{2m+1}^{(b)}$, we obtain the same rate of convergence $b^{-(m+1)}$ with different leading constants. This is due to the moment properties of $T_b(s)$ stated in the following.

Lemma 6.1. *Let $b \geqslant 1$, $s > 0$ and $m = 2, 3, \ldots$. If m is odd (even), then*

$$\mu_m(b, s) := \mathrm{E}(T_b(s) - s)^m = \left(\frac{s}{b}\right)^{m/2}\left(EZ^m + \sum_{\substack{l=1 \\ l:\ \text{odd (even)}}}^{m-2} \alpha(l)(bs)^{-l/2}\right) \geqslant 0,$$

where Z is a standard normal random variable, $\alpha(l)$, $l = 1, \ldots, m-2$, is a positive constant not depending on b or s, and $\sum_1^0 := 0$.

Lemma 6.1 may be derived from Von Bahr (1965, Theorems 1 and 3), except for two minor details. First, Von Bahr's results are stated for normalized sums of independent random variables in a discrete time setting. Second, it is not clear from Von Bahr's paper if the coefficients $\alpha(l)$ are positive. In its actual form, Lemma 6.1 is proved by Adell (2004).

To give a closed-form expression for the constants appearing in the main terms of the approximation, we introduce, for any $b \geqslant 1$ and $s > 0$, the sequence of polynomials $(Q_k(b, s; x),\ k \in \mathbb{N})$ defined by

$$\begin{aligned} Q_{2m}(b, s; x) \quad &:= \frac{b^{m+1}}{2}\left(\frac{\mu_{2m+2}(b, s)}{(2m+2)!}C_{2m+2}(s; x) - \frac{\mu_{2m+1}(b, s)}{(2m+1)!}C_{2m+1}(s; x)\right), \\ Q_{2m+1}(b, s; x) \quad &:= \frac{b^{m+1}}{2}\frac{\mu_{2m+2}(b, s)}{(2m+2)!}C_{2m+2}(s; x), \qquad x \geqslant 0, \quad m \in \mathbb{N}. \end{aligned} \tag{6.1}$$

Henceforth, C stands for a positive constant independent of b and s whose value may change from line to line. We shall need the following.

Lemma 6.2. *For any $k \in \mathbb{N}$, we have*

$$\sup_{b\geqslant 1, s>0} \mathrm{E}|Q_k(b,\, s;\, N(s))| \leqslant C. \tag{6.2}$$

In addition, we have

$$\sup_{b\geqslant 1, s>0} \mathrm{E}|Q_0(b,\, s;\, N(s))| = \frac{1}{4}\sup_{s>0} s\mathrm{E}|C_2(s;\, N(s))| = \frac{3}{4\mathrm{e}}. \tag{6.3}$$

Proof. By Lemma 6.1, inequality (6.2) will follow once we show that

$$\mathrm{E}|C_m(s;\, N(s))| \leqslant C(1 \wedge s^{-m/2}), \qquad s > 0,\ m \in \mathbb{N}. \tag{6.4}$$

Indeed, from (2.2) we have by calculus that $\mathrm{E}|C_m(s;\, N(s))| \leqslant C,\ \ 0 < s \leqslant 1,\ \ m \in \mathbb{N}$. Therefore, (6.4) follows from (2.3), thus completing the proof of inequality (6.2).

By considering the roots of $C_2(s; x)$, it can be seen by calculus that $sE|C_2(s;\, N(s))| \leqslant \mathrm{E}|C_2(1,\, N(1))| = 3\mathrm{e}^{-1},\ \ 0 < s \leqslant 1$. On the other hand, Roos (2003a, Lemma 4) has shown that $sE|C_2(s;\, N(s))| \leqslant 3\mathrm{e}^{-1},\ 1 \leqslant s$. This shows (6.3) and completes the proof of the lemma. □

Keeping in mind (6.1), we have the following.

Theorem 6.1. *Let $b \geqslant 1,\ s > 0$ and $m \in \mathbb{N}$. Then,*

$$\left| d_{\mathrm{TV}}(\mathcal{P}(T_b(s)),\, \nu_{2m}^{(b)}) - b^{-(m+1)}\,\mathrm{E}|Q_{2m}(b,\, s;\, N(s))| \right| \leqslant Cb^{-(m+2)} \tag{6.5}$$

and

$$\left| d_{\mathrm{TV}}(\mathcal{P}(T_b(s)),\, \nu_{2m+1}^{(b)}) - b^{-(m+1)}\,\mathrm{E}|Q_{2m+1}(b,\, s;\, N(s))| \right| \leqslant Cb^{-(m+2)}. \tag{6.6}$$

In particular, setting $m = 0$ in (6.5) *or* (6.6)*, we have*

$$\left| d_{\mathrm{TV}}(\mathcal{P}(T_b(s)),\, \mathcal{P}(s)) - \frac{s}{4b}\,\mathrm{E}|C_2(s;\, N(s))| \right| \leqslant \frac{C}{b^2}. \tag{6.7}$$

Proof. The proof of (6.6) being similar, we shall only show (6.5). Let $b \geqslant 1$ and $s > 0$. Let $\phi \in \mathcal{E}(\alpha)$ be such that $\|\Delta_{2m+4}\phi\| < \infty$. By (2.5) and Corollary 2.1, we can write

$$\begin{aligned} &\mathrm{E}\phi(N(T_b(s))) - \int_{\mathbb{N}} \phi(l)\nu_{2m}^{(b)}(\mathrm{d}l) \\ &\quad = 2b^{-(m+1)}\mathrm{E}\phi(N(s))Q_{2m}(b,\, s;\, N(s)) - \frac{\mu_{2m+3}(b,\, s)}{(2m+3)!}\,\mathrm{E}\phi(N(s))C_{2m+3}(s;\, N(s)) \\ &\qquad + \frac{1}{(2m+4)!}\mathrm{E}(T_b(s) - s)^{2m+4}\Delta_{2m+4}\phi(N(s + (T_b(s) - s)\beta_{2m+4})). \end{aligned} \tag{6.8}$$

Concerning the main term in (6.8), we have as in (3.2) that

$$\sup_{A\subseteq\mathbb{N}}|\mathrm{E}1_A(N(s))Q_{2m}(b,\, s;\, N(s))| = \tfrac{1}{2}\mathrm{E}|Q_{2m}(b,\, s;\, N(s))|. \tag{6.9}$$

By (3.2), Lemma 6.1 and (6.4), it can be verified that

$$\sup_{A\subseteq\mathbb{N}} \mu_{2m+3}(b,\, s)|\mathrm{E}1_A(N(s))C_{2m+3}(s;\, N(s))| \leqslant Cb^{-(m+2)}. \tag{6.10}$$

Finally, by Lemma 2.1 and (6.4), we have

$$\sup_{A\subseteq\mathbb{N}}|\mathrm{E}(T_b(s)-s)^{2m+4}\Delta_{2m+4}1_A(N(s+(T_b(s)-s)\beta_{2m+4}))|$$

$$\leqslant C\min\left\{\mathrm{E}(T_b(s)-s)^{2m+4},\, \mathrm{E}\left(\frac{T_b(s)-s}{\sqrt{s+(T_b(s)-s)\beta_{2m+4}}}\right)^{2m+4}\right\}. \tag{6.11}$$

Choose a constant τ large enough so that if $bs \geqslant \tau$ the second expectation on the right-hand side in (6.11) exists. Then,

$$\sup_{0<bs\leqslant\tau} \mathrm{E}(T_b(s)-s)^{2m+4} = \frac{1}{b^{2m+4}}\sup_{0<t\leqslant\tau} \mathrm{E}(X(t)-t)^{2m+4} \leqslant \frac{C}{b^{2m+4}} \tag{6.12}$$

and

$$\sup_{bs\geqslant\tau} \mathrm{E}\left(\frac{T_b(s)-s}{\sqrt{s+(T_b(s)-s)\beta_{2m+4}}}\right)^{2m+4} \leqslant \frac{C}{b^{m+2}}, \tag{6.13}$$

where the last inequality follows from the central limit theorem and the strong law of large numbers for the gamma process. As in the proof of Theorem 3.1(b), (6.5) follows from (6.8)–(6.13). □

Many authors have obtained estimates in total variation between negative binomial and Poisson distributions. For instance, Barbour (1987, p. 758) has shown that

$$d_{\mathrm{TV}}(\mathcal{P}(T_b(s)),\, \mathcal{P}(s)) \leqslant \frac{1-\mathrm{e}^{-s}}{b} \leqslant \frac{s\wedge 1}{b} \tag{6.14}$$

and Roos (2003a, formula (31)) has given the estimate

$$d_{\mathrm{TV}}(\mathcal{P}(T_b(s)),\, \mathcal{P}(s)) \leqslant \frac{1}{b}\left(s\wedge\frac{3}{4\mathrm{e}}\right). \tag{6.15}$$

For other upper bounds, we refer to Roos (2003a, formulae (32)–(35)). By calculus, we have $\mathrm{E}|C_2(s;\, N(s))| = 2(2-s)\mathrm{e}^{-s}$, $0 < s < 2-\sqrt{2}$. Therefore, by (6.7), estimate (6.14) is sharp as $s \to 0$. Also, (6.3) and (6.7) show that the constant $3/4\mathrm{e}$ in (6.15) is sharp (this fact has already been noted by Roos (2003a, Theorem 1) in a more general setting). Finally, using the central limit theorem and the strong law of large numbers for the Poisson process, it is not hard to see that $sE|C_2(s;\, N(s))| \to (8/(\pi\mathrm{e}))^{-1/2}$ as $s \to \infty$. In this sense, Roos (2003a, formula (38)) has obtained the asymptotic relation

$$|b\sqrt{2\pi e}\, d_{\mathrm{TV}}(\mathcal{P}(T_b(s)), \mathcal{P}(s)) - 1| \leqslant C\left(\frac{1}{s} + \frac{1}{b\sqrt{s}} + \frac{1}{b}\right). \tag{6.16}$$

With the help of Lemma 5 in Roos (2003a), a similar asymptotic relation to that in (6.16) may be derived from (6.7). Details are omitted.

Acknowledgements

We are very grateful to the referees for their careful reading of the manuscript and their constructive criticism. Thanks are also due to one of the referees for having pointed out some inaccuracies and errors in the previous version, together with their solution.

This work was supported by research grants no. BFM2002-04163-C02-01 and no. DGA E-12/25.

References

Adell, J.A. (2004) A differential calculus for linear operators and its application to the central limit theorem. Preprint.

Adell, J.A. and Lekuona, A. (2000) Taylor's formula and preservation of generalized convexity for positive linear operators. *J. Appl. Probab.*, **37**, 765–777.

Adell, J.A. and Sangüesa, C. (2004) Approximation by B-spline convolution operators. A probabilistic approach. To appear in *J. Comput. Appl. Math*.

Barbour, A.D. (1987) Asymptotic expansions in the Poisson limit theorem. *Ann. Probab.*, **15**, 748–766.

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

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

Čekanavičius, V. and Kruopis, J. (2000) Signed Poisson approximation: a possible alternative to normal and Poisson laws. *Bernoulli*, **6**, 591–606.

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

Chihara, T.S. (1978) *An Introduction to Orthogonal Polynomials*. New York: Gordon and Breach.

Deheuvels, P. and Pfeifer, D. (1986) A semigroup approach to Poisson approximation. *Ann. Probab.*, **14**, 663–676.

Deheuvels, P., Pfeifer, D. and Puri, M.L. (1989) A new semigroup technique in Poisson approximation. *Semigroup Forum*, **38**, 189–201.

Denuit, M. and Van Bellegem, S. (2001) On the stop-loss and total variation distances between random sums. *Statist. Probab. Lett.*, **53**, 153–165.

Grandell, J. (1997) *Mixed Poisson Processes*. London: Chapman & Hall.

Johnson, N.L., Kotz, S. and Kemp, A.W. (1992) *Univariate Discrete Distributions*, 2nd Edition. New York: Wiley.

Kornya, P.S. (1983) Distribution of aggregate claims in the individual risk theory model. *Trans. Soc. Actuaries*, **35**, 823–858.

Pfeifer, D. (1987) On the distance between mixed Poisson and Poisson distributions. *Statist. Decisions*, **5**, 367—379.

Presman, É.L. (1983) Approximation of binomial distributions by infinitely divisible ones. *Theory Probab. Appl.*, **28**, 393–403.

Rachev, S.T. (1991) *Probability Metrics and the Stability of Stochastic Models*. Chichester: Wiley.

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.

Roos, B. (1999) Asymptotics and sharp bounds in the Poisson approximation to the Poisson-binomial distribution. *Bernoulli*, **5**, 1021–1034.

Roos, B. (2001) Sharp constants in the Poisson approximation. *Statist. Probab. Lett.*, **52**, 155–168.

Roos, B. (2002) Kerstan's method in the multivariate Poisson approximation: an expansion in the exponent. *Theory Probab. Appl.*, **47**, 358–363.

Roos, B. (2003a) Improvements in the Poisson approximation of mixed Poisson distributions. *J. Statist. Plann. Inference*, **113**(2), 467–483.

Roos, B. (2003b) Poisson approximation of multivariate Poisson mixtures. *J. Appl. Probab.*, **40**, 376–390.

Serfling, R. J. (1978) Some elementary results on Poisson approximation in a sequence of Bernoulli trials. *SIAM Rev.*, **20**, 567–579.

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.

Von Bahr, B. (1965) On the convergence of moments in the central limit theorem. *Ann. Math. Statist.*, **36**, 808–818.

Received November 2002 and revised October 2003