UNIVERSITÄT HAMBURG

Improvements in the Poisson approximation of mixed Poisson distributions

Bero Roos

Preprint–No. 2001–06 November 2001

FACHBEREICH MATHEMATIK

SCHWERPUNKT MATHEMATISCHE STATISTIK
UND STOCHASTISCHE PROZESSE

BUNDESSTR. 55, D – 20146 HAMBURG

IMPROVEMENTS IN THE POISSON APPROXIMATION OF MIXED POISSON DISTRIBUTIONS

Bero Roos

Schwerpunkt Math. Statistik und Stoch. Prozesse, Fachbereich Mathematik, Universität Hamburg, Bundesstraße 55, D–20146 Hamburg, Germany.

E–mail: roos@math.uni-hamburg.de

Abstract. We consider the approximation of mixed Poisson distributions by Poisson laws and also by related finite signed measures of higher order. Upper bounds and asymptotic relations are given for several distances. Even in the case of the Poisson approximation with respect to the total variation distance, our bounds have better order than those given in the literature. In particular, our results hold under weaker moment conditions for the mixing random variable. As an example, we consider the approximation of the negative binomial distribution, which enables us to prove the sharpness of a constant in the upper bound of the total variation distance. The main tool is an integral formula for the difference of the counting densities of a Poisson distribution and a related finite signed measure.

MSC 2000. Primary 60F05; secondary 62E17.

Key words and phrases. Asymptotic relations, finite signed measures, mixed Poisson distribution, negative binomial distribution, Poisson approximation, probability metrics, sharp constant, total variation distance, weak moment conditions.

1 Introduction

Mixed Poisson distributions are widely used in probability theory and statistics. In the books by Haight (1967), Douglas (1980), Johnson et al. (1993), and Grandell (1997), one can find an overview of applications and general properties. A large list of examples of mixed Poisson distributions can be found in Johnson et al. (1993, p. 328–335). For instance the following distributions are mixed Poisson: the negative binomial distribution (see also Section 4), the Delaporte distribution in actuarial sciences, the truncated-gamma mixture of Poisson distributions in the context of limited collective risk theory, the Neyman Type A distribution in biology and ecology, the Poisson-Pascal and Pólya-Aeppli distributions in biology, the Poisson-Lindley distribution with applications to errors and accidents.

Mixed Poisson distributions can be very involved. This is the case when the distribution of the mixing random variable has a complicated structure. For example, the convolution of mixed Poisson distributions is again mixed

Poisson with mixing random variable $X = \sum_{i=1}^{n} X_i$, where we suppose that the $X_1, \ldots, X_n$ are independent mixing random variables of the corresponding components (see Feller (1943) or Johnson et al. (1993, p. 327)). Hence, it is often necessary to use an approximation of mixed Poisson distributions. If the mixing random variable is almost constant, it is reasonable to apply a simple Poisson distribution.

In the present paper, we consider the approximation of mixed Poisson distributions by Poisson distributions and also by related finite signed measures of higher order. We present upper bounds and asymptotic relations for several distances. The main tool is an integral formula for the difference of the counting densities of a Poisson distribution and a related finite signed measure. Before we treat the more general case (see Section 3), we shall discuss our results in the case of Poisson approximation concerning the total variation distance.

Let G denote a mixed Poisson distribution with the counting density

$$g(m) = G(\{m\}) = \mathrm{E}(\pi(m, X)), \qquad (m \in \mathbf{Z}_+ = \{0, 1, 2, \ldots\}), \tag{1}$$

where E means expectation, X is the mixing random variable concentrated on $[0, \infty)$, and $\pi(m, t) = e^{-t} t^m / m!$, $(m \in \mathbf{Z}_+,\ t \in [0, \infty))$ denotes the counting density of the Poisson distribution Π_t with mean t. Here and throughout this paper, we set $0^0 = 1$, such that $\Pi_0 = \varepsilon_0$ is the Dirac measure at point 0.

In Pfeifer (1987) and Barbour et al. (1992, pp. 12–13 and 68–69), one can find some results concerning the Poisson approximation of mixed Poisson distributions (see also Remarks 2(b) and 4 of the present paper). In particular, there are some error bounds for the total variation distance d_{TV} (for a definition see Subsection 2.3). The most remarkable inequalities came from Barbour et al. (1992, Theorem 1.C, Remark 1.1.2), who used the Stein–Chen method to prove that, letting $\sigma^2 = \mathrm{Var}(X)$,

$$d_{TV}(G, \Pi_\mu) \ \le\ (1 - e^{-\mu})\,\frac{\sigma^2}{\mu} \le \sigma^2 \min\left\{\frac{1}{\mu}, 1\right\}, \quad \text{if } \mu = \mathrm{E}X \in (0, \infty), \tag{2}$$

$$d_{TV}(G, \Pi_s) \ \le\ \mathrm{E}|X - s| \min\left\{\sqrt{\frac{2}{se}}, 1\right\}, \qquad \text{if } s \in (0, \infty). \tag{3}$$

In the majority of cases, a Poisson distribution with mean μ will be used as an approximation, so that, for applications, (2) is more important than (3). According to (2), Poisson approximation is good, if the variance σ^2 or the quotient σ^2/μ is small. Since in the Poisson approximation of the Poisson binomial distribution the Stein–Chen method led to upper bounds of the correct order with sharp constants (see Barbour et al. (1992, Corollary 3.D.1)), one may ask, whether the estimates in (2) and (3) are best possible. The results of this paper imply that the answer is no. In particular, we have proved the following theorem.

Theorem 1 (a) *The following estimates are valid:*

$$d_{TV}(G,\,\Pi_\mu) \;\leq\; \min\Big\{\frac{3}{2e}\mathrm{E}\Big(X\ln\frac{X}{\mu}\Big),\,\sigma^2\Big\}, \qquad \textit{if } \mu\in(0,\infty), \tag{4}$$

$$d_{TV}(G,\,\Pi_s) \;\leq\; \min\Big\{\sqrt{\frac{2}{e}}\mathrm{E}|\sqrt{X}-\sqrt{s}|,\,\mathrm{E}|X-s|\Big\}, \quad \textit{if } s\in[0,\infty). \tag{5}$$

(b) *The constant* $3/(2e)$ *in* (4) *cannot be reduced.*

Here and throughout this paper, we set $0\ln 0 = 0$. Part (a) of the theorem is a direct consequence of Theorem 2 below. For a proof, see Section 5. Further, we show part (b) of the theorem by considering a suitable example (see Section 4). For the comparison of (2) and (3) with the bounds in Theorem 1, we use the following elementary lemma, the proof of which is located in Section 5 as well.

Lemma 1 *Let* $\mu_r = \mathrm{E}(X^r)$ *for* $r\in[0,\infty)$. *If* $\mu=\mu_1\in(0,\infty)$ *and* $\epsilon\in(0,\infty)$, *then*

$$\mathrm{E}\Big(X\,\ln\frac{X}{\mu}\Big) \leq \frac{\mu}{\epsilon}\,\ln\frac{\mu_{1+\epsilon}}{\mu^{1+\epsilon}} \leq \frac{1}{\epsilon\,\mu^\epsilon}\,\Big(\mu_{1+\epsilon}-\mu^{1+\epsilon}\Big). \tag{6}$$

If $\epsilon\in[0,1/2]$ *and* $s\in(0,\infty)$, *then*

$$\mathrm{E}|\sqrt{X}-\sqrt{s}| \leq \frac{\mathrm{E}|X-s|^{1/2+\epsilon}}{s^\epsilon}. \tag{7}$$

From (6), we obtain in the case $\epsilon=1$ that $\mathrm{E}(X\,\ln(X/\mu)) \leq \mu\ln(\mu_2/\mu^2) \leq \sigma^2/\mu$, if X has finite variance σ^2. Further, letting $\epsilon = 1/2$ in (7), we see that, for $s\in(0,\infty)$, $\mathrm{E}|\sqrt{X}-\sqrt{s}| \leq \mathrm{E}|X-s|/\sqrt{s}$. Therefore (4) and (5) are substantially better than the second inequality in (2) and (3) obtained by Barbour et al. (1992). It is noteworthy that, in comparison with (2) and (3), our bounds hold under weaker moment conditions. For example, in the case of infinite variance σ^2 and finite moment $\mu_{1+\epsilon}$, the bounds in (2) are infinite, whereas our bound (4) remains finite. A similar example is possible for the comparison of (3) with (5). Note that, by (5), we obtain an upper bound for the total variation distance between two Poisson distributions with different means:

$$d_{TV}(\Pi_t,\,\Pi_s) \leq \min\Big\{\sqrt{\frac{2}{e}}\,|\sqrt{t}-\sqrt{s}|,\,|t-s|\Big\}, \qquad (s,t\in[0,\infty)).$$

Parts of this bound are due to Freedman (1974, (8)) and Daley and Vere–Jones (1988, p. 300). For a cruder bound, see Yannaros (1991, Theorem 2.1).

The paper is structured as follows. In the next section, we collect necessary notation about moments, finite signed measures, and probability metrics. Section 3 is devoted to the main results. Here we present upper bounds and asymptotic relations. In Section 4, we discuss the negative binomial distribution as an example of a mixed Poisson distribution and give a proof of Theorem 1(b). In Section 5, we present the remaining proofs.

2 Further notation

2.1 Moments

As above, let $\mu_r = \mathrm{E}(X^r)$, $(r \in [0,\infty))$ be the r-th moment of X. We set $\mu = \mu_1$ and always assume that $\mu > 0$. Further, the j-th moment (resp. the r-th absolute moment) of X about $s \in [0,\infty)$, if it exists, is given by $\nu'_{j,s} = \mathrm{E}(X-s)^j$ (resp. $\nu_{r,s} = \mathrm{E}|X-s|^r$), where $j \in \mathbf{Z}_+$ and $r \in [0,\infty)$. Therefore, if it exists, $\sigma^2 = \nu'_{2,\mu} = \nu_{2,\mu}$ is the variance of X.

2.2 Finite signed measures

As candidates for an approximation, we choose a Poisson distribution Π_s with mean $s \in [0,\infty)$ and the related finite signed measures $G_{k,s}$, $(k \in \mathbf{Z}_+,\ s \in [0,\infty))$ with the counting density

$$g_{k,s}(m) = G_{k,s}(\{m\}) = \sum_{j=0}^{k} \frac{\nu'_{j,s}}{j!}\,\Delta^j \pi(m,s), \qquad (m \in \mathbf{Z}_+). \tag{8}$$

Here $\Delta^j = \Delta \circ \ldots \circ \Delta$, $(j \in \mathbf{N} = \{1,2,3,\ldots\})$ is the j-th iterated composition of the difference operator Δ from $\mathbf{R}^{\mathbf{Z}_+} = \{f \mid f : \mathbf{Z}_+ \longrightarrow \mathbf{R}\}$ onto itself, defined by $\Delta f(m) = f(m-1) - f(m)$ for $f \in \mathbf{R}^{\mathbf{Z}_+}$ and $m \in \mathbf{Z}_+$, with the convention that $f(m) = 0$ for all $m < 0$. Further Δ^0 is the identity of $\mathbf{R}^{\mathbf{Z}_+}$ onto itself. In (8) and henceforth, we use the notation $\Delta^j \pi(m,s) = (\Delta^j \pi(\cdot,s))(m)$. For the inverse operator $\Delta^{-1} : \mathbf{R}^{\mathbf{Z}_+} \longrightarrow \mathbf{R}^{\mathbf{Z}_+}$, we have $\Delta^{-1} f(m) = -\sum_{k=0}^{m} f(k)$, $(f \in \mathbf{R}^{\mathbf{Z}_+},\ m \in \mathbf{Z}_+)$. Similar to above, let Δ^{-j} be the j-th iterated composition of Δ^{-1}.

Remark 1 (a) If we use $G_{k,s}$ with $k \in \mathbf{N}$ for an approximation of G, we must assume that $\mu_k < \infty$.

(b) Let us collect some facts about the $G_{k,s}$: First note that, for each $s \in [0,\infty)$, we have $G_{0,s} = \Pi_s$. Further, the counting density of $G_{k,s}$ for $k \in \mathbf{N}$ and $s \in [0,\infty)$ can be evaluated by the help of the following well-known formula (see, for example, Roos (1999, formula (6))):

$$\Delta^k \pi(m,s) = \frac{1}{s^k}\,\pi(m,s)\,C_k(m,s), \qquad (s \in (0,\infty), k, m \in \mathbf{Z}_+), \tag{9}$$

where

$$C_k(x,s) = \sum_{j=0}^{k} \binom{k}{j}\binom{x}{j}\, j!\,(-s)^{k-j}, \qquad (s, x \in \mathbf{R},\ k \in \mathbf{Z}_+)$$

denotes the Charlier polynomial of degree k. We obtain, for $s \in (0,\infty)$ and $m \in \mathbf{Z}_+$,

$$\begin{aligned} g_{1,s}(m) &= \pi(m,s)\Big(1+\frac{(\mu-s)(m-s)}{s}\Big),\\ g_{2,s}(m) &= \pi(m,s)\Big(1+\frac{(\mu-s)(m-s)}{s}+\frac{\nu'_{2,s}}{2s^2}(m^2-(2s+1)m+s^2)\Big). \end{aligned}$$

The case $s=0$ can be treated by letting $s \downarrow 0$ in the formulas above. It should be mentioned that, if $\mu<\infty$, we have $G_{1,\mu}=\Pi_\mu$.

2.3 Probability metrics

The accuracy of approximation will be measured by the distances

$$d_p^{(i)}(H_1,H_2) = \|\Delta^i(h_1-h_2)\|_p, \tag{10}$$

where H_1 and H_2 are two finite signed measures concentrated on $\mathbf{Z}_+$ with counting densities $h_1, h_2 \in \mathbf{R}^{\mathbf{Z}_+}$. Here, we set $\|f\|_1 = \sum_{m=0}^\infty |f(m)|$ and $\|f\|_\infty = \sup_{m\in\mathbf{Z}_+}|f(m)|$ for $f \in \mathbf{R}^{\mathbf{Z}_+}$. Further, we assume that $p \in \{1,\infty\}$ and $i \in \mathbf{Z}$. If G is approximated by $G_{k,s}$, ($k\in\mathbf{Z}_+$, $s\in[0,\infty)$), we additionally assume in Theorem 2 below that $k+i+1\geq 0$. In this case, we have to take $h_1=g$ and $h_2=g_{k,s}$ (see (1) and (8)). The definition (10) leads to several well-known probability metrics, if we consider probability distributions H_1 and H_2: We obtain the total variation distance $d_{TV}=d_1^{(0)}/2$, the Kolmogorov metric $d_\infty^{(-1)}$, the Fortet–Mourier metric $d_1^{(-1)}$, the point metric $d_\infty^{(0)}$, and the stop–loss distances $d_\infty^{(-2)}$ and $d_1^{(-2)}$. For the general theory of probability metrics, we refer the reader to Rachev (1991).

3 Main results

3.1 Upper bounds

The following theorem forms the main result of this paper. Here and throughout the paper, let $\lfloor x\rfloor \in \mathbf{Z}$ be the largest integer $\leq x \in \mathbf{R}$. Set, as usual, $1/\infty = 0$.

Theorem 2 *Let $p\in\{1,\infty\}$, $s\in[0,\infty)$, $k\in\mathbf{Z}_+$, and $i\in\mathbf{Z}$ with $k+i+1\geq 0$. Then*

$$\left.\begin{aligned} d_1^{(i)}(G,\,G_{k,s}) &\leq \frac{2^{k+i+1}}{(k+1)!}\,\nu_{k+1,s},\\ d_\infty^{(i)}(G,\,G_{k,s}) &\leq \binom{k+i+1}{\lfloor (k+i+1)/2\rfloor}\frac{\nu_{k+1,s}}{(k+1)!}. \end{aligned}\right\} \tag{11}$$

Further

$$d_p^{(i)}(G,\,G_{k,s}) \leq \frac{U_p^{(k+i+1)}}{k!}\,\mathrm{E}\left|\int_s^X \frac{|X-y|^k\,dy}{y^{(k+i+2)/2-1/(2p)}}\right|. \tag{12}$$

where the constants $U_p^{(n)}$, $(p \in \{1, \infty\},\ n \in \mathbf{Z}_+)$ *are defined in Lemma 4 below and satisfy*

$$\left.\begin{array}{lll} U_1^{(0)} = 1, & U_1^{(1)} = \sqrt{\dfrac{2}{e}}, & U_1^{(2)} = \dfrac{3}{e}, \\ U_\infty^{(0)} = \dfrac{1}{\sqrt{2e}}, & U_\infty^{(1)} = \dfrac{1}{e}, & U_\infty^{(2)} = \left(\dfrac{3}{2e}\right)^{3/2}, \end{array}\right\} \tag{13}$$

and

$$U_1^{(n)} \le \sqrt{n!}, \qquad U_\infty^{(n)} \le \frac{\sqrt{e}}{2}\left(1 + \sqrt{\frac{\pi}{2n}}\right)\left(\frac{n}{e}\right)^{(n+1)/2}, \qquad (n \in \mathbf{N}). \tag{14}$$

Remark 2 (a) In (12) and throughout the paper, we use the equality "$\int_a^b = -\int_b^a$". Therefore it can occur that the integral in (12) is negative. Sometimes the following easy equality is useful: For $a, s \in [0, \infty)$,

$$\left| \int_s^X \frac{|X-y|^k}{y^a}\, dy \right| = |X - s|^{k+1} \int_0^1 \frac{(1-y)^k\, dy}{(yX + (1-y)s)^a}. \tag{15}$$

(b) The inequalities in (11) are generalizations of (2.1) in Pfeifer (1987).

The bounds in Theorem 1 (see Section 1) and in the next corollary are easy consequences of Theorem 2. Here, we only consider the unbiased case, where $s = \mu$ is finite. It is easy to derive similar results for the biased case $s \ne \mu$.

Corollary 1 *If* $\mu \in (0, \infty)$, *the following relations are valid under the assumption that the right–hand sides exist:*

$$d_1^{(-2)}(G, \Pi_\mu) \;=\; \frac{\sigma^2}{2}, \tag{16}$$

$$d_\infty^{(-2)}(G, \Pi_\mu) \;\le\; \min\left\{\frac{4}{3\sqrt{2e}}\,(\mu_{3/2} - \mu^{3/2}),\ \frac{\sigma^2}{2}\right\}, \tag{17}$$

$$d_1^{(-1)}(G, \Pi_\mu) \;\le\; \min\left\{\frac{4}{3}\sqrt{\frac{2}{e}}\,(\mu_{3/2} - \mu^{3/2}),\ \sigma^2\right\}, \tag{18}$$

$$d_\infty^{(-1)}(G, \Pi_\mu) \;\le\; \min\left\{\frac{1}{e}\mathrm{E}\Big(X \ln\frac{X}{\mu}\Big),\ \frac{\sigma^2}{2}\right\}, \tag{19}$$

$$d_\infty^{(0)}(G, \Pi_\mu) \;\le\; \min\left\{4\Big(\frac{3}{2e}\Big)^{3/2}(\mu^{1/2} - \mu_{1/2}),\ \sigma^2\right\}. \tag{20}$$

Note that, in (16), indeed equality holds. Further, we have no doubt that, similar to the statement (b) in Theorem 1, the corresponding constants in the bounds of Corollary 1 are sharp.

3.2 Asymptotic relations

Sharper inequalities than those given in Theorem 2 and asymptotic relations are possible by means of

$$\left| d_p^{(i)}(G,\, G_{k,s}) - \frac{|\nu'_{k+1,s}|}{(k+1)!}\|\Delta^{k+i+1}\pi(\cdot, s)\|_p \right| \le d_p^{(i)}(G,\, G_{k+1,s}) \tag{21}$$

($p \in \{1, \infty\}$, $s \in [0, \infty)$, $k \in \mathbf{Z}_+$, $i \in \mathbf{Z}$ with $k+i+1 \geq 0$) and the inequalities in Theorem 2, whereby we must know something about the norm $\|\Delta^{k+i+1}\pi(\cdot, s)\|_p$. In Roos (1999, Sections 3 and 4), one can find some properties of such norm terms. Some special cases were investigated by Deheuvels and Pfeifer (1986a, b). For sharp upper bounds and asymptotic relations of the norms, we refer the reader to Lemmas 3–5.

In what follows, we only consider results concerning the total variation distance in the unbiased case. The other distances and the biased case can be treated similarly.

Theorem 3 *Let us assume that* $\mu \in (0, \infty)$*. For* $t \in [0, \infty)$ *and* $m \in \mathbf{Z}_+$*, set*

$$L_m(t) = (m-t)\,\pi(m,t), \qquad b_\pm(t) = \Big\lfloor t + \frac{1}{2} \pm \sqrt{t + \frac{1}{4}} \Big\rfloor.$$

(a) *The relations*

$$d_{TV}(G,\,\Pi_\mu) = \frac{\sigma^2}{2\mu}\Big(L_{b_+(\mu)}(\mu) - L_{b_-(\mu)}(\mu)\Big) + R \leq \frac{3\sigma^2}{4\mu e} + R \tag{22}$$

hold, where $|R| \leq d_{TV}(G,\,G_{2,\mu})$*,*

$$d_{TV}(G,\,G_{2,\mu}) \leq \min\Big\{\frac{U_1^{(3)}}{2\sqrt{\mu}}\mathrm{E}\Big(\frac{|X-\mu|^3}{(\sqrt{X}+\sqrt{\mu})^2}\Big),\, \frac{2}{3}\,\nu_{3,\mu}\Big\}, \tag{23}$$

$$d_{TV}(G,\,G_{2,\mu}) \leq \min\Big\{\frac{U_1^{(3)}\,|\nu'_{3,\mu}|}{12\,\mu^{3/2}} + \frac{U_1^{(4)}}{12\,\mu}\mathrm{E}\Big(\frac{(X-\mu)^4}{X+2\mu}\Big),\, \frac{1}{3}\,(2|\nu'_{3,\mu}| + \nu_{4,\mu})\Big\}. \tag{24}$$

In (22)*, equalities hold, if* $\mu = 1$.

(b) *An absolute constant* $c_1 \in (0, \infty)$ *exists such that*

$$\Big|\frac{d_{TV}(G,\,\Pi_\mu)}{\sigma^2/(\mu\sqrt{2\pi e})} - 1\Big| \leq c_1\Big(\frac{1}{\mu} + \frac{\mu}{\sigma^2}\,d_{TV}(G,\,G_{2,\mu})\Big). \tag{25}$$

Remark 3 (a) By Theorem 3, several asymptotic relations for $d_{TV}(G, \Pi_\mu)$ can be proved. In particular, from Theorem 3(b) it follows that $d_{TV}(G,\,\Pi_\mu) \sim \sigma^2/(\mu\sqrt{2\pi e})$, whenever $\mu \to \infty$ and $d_{TV}(G,\,G_{2,\mu})\,\mu/\sigma^2 \to 0$. Here we set $a_n \sim b_n$, if $a_n/b_n \to 1$. For an application concerning the negative binomial distribution, see Section 4. Further, we need Theorem 3 for the proof of Theorem 1(b).

(b) Suitable upper bounds of the right–hand sides of (23) and (24) can be derived by means of the easy inequalities

$$\mathrm{E}\Big(\frac{|X-s|^3}{(\sqrt{X}+\sqrt{s})^2}\Big) \leq \frac{\nu_{2+\epsilon,s}}{s^\epsilon}, \qquad \mathrm{E}\Big(\frac{(X-s)^4}{X+2s}\Big) \leq \frac{\nu_{3+\epsilon,s}}{(2s)^\epsilon} \tag{26}$$

for $s \in (0, \infty)$ and $\epsilon \in [0, 1]$. In practice, (24) has sometimes advantages over (23), since here we do not have to deal with absolute moments of third order. An application of (24) is given below (see (37)).

(c) For an upper bound of $d_{TV}(G,\,G_{3,\mu})$, consult (45).

Remark 4 By using the semigroup method, Pfeifer (1987, (2.20)) proved the following asymptotic relation: If the distribution of X depends on the mean μ and the variance σ^2 in such a way that

$$\nu_{4,\mu} = O\Big(\frac{\sigma^2}{\mu}\Big), \qquad \mu \to \infty, \tag{27}$$

then

$$d_{TV}(G, \Pi_\mu) = \frac{\sigma^2}{\mu\sqrt{2\pi e}}\Big(1 + O\Big(\frac{1}{\mu}\Big)\Big), \tag{28}$$

where $f_1(x) = O(f_2(x))$ means that $|f_1(x)/f_2(x)|$ is bounded. For a comparison with our results, we look at (23) and (25) and see that (28) is valid, if we suppose the condition

$$\mathrm{E}\Big(\frac{|X-\mu|^3}{(\sqrt{X}+\sqrt{\mu})^2}\Big) = O\Big(\frac{\sigma^2}{\mu^{3/2}}\Big). \tag{29}$$

Here we do not need the assumption $\mu \to \infty$. From the Hölder inequality it easily follows that condition (29) is weaker than (27). Indeed, if (27) holds, we obtain

$$\mathrm{E}\Big(\frac{|X-\mu|^3}{(\sqrt{X}+\sqrt{\mu})^2}\Big) \le \frac{\nu_{3,\mu}}{\mu} \le \frac{\sigma\sqrt{\nu_{4,\mu}}}{\mu} = O\Big(\frac{\sigma^2}{\mu^{3/2}}\Big).$$

Hence, our Theorem 3 is stronger than Pfeifer's result.

4 An example

4.1 Motivation

As an important example of a mixed Poisson distribution, we discuss in this section the negative binomial distribution, for which several bounds are already available in the literature. A comparison with our results document how good our results are. Finally we are concerned with asymptotic relations, which lead to a proof of the sharpness result (b) in Theorem 1.

The negative binomial distribution $G = \mathrm{NB}(q,\alpha)$ with parameters $q \in (0,1)$ and $\alpha \in (0,\infty)$, is given by

$$\mathrm{NB}(q,\alpha)(\{m\}) = \binom{\alpha+m-1}{m}(1-q)^\alpha\, q^m, \qquad (m \in \mathbf{Z}_+).$$

Here, the mixing random variable X has a gamma distribution with Lebesgue–density

$$\gamma(t) = \frac{\beta^\alpha}{\Gamma(\alpha)}\, t^{\alpha-1}\, e^{-t\beta}, \qquad (t \in (0,\infty)),$$

where $\beta = (1-q)/q$ (see Johnson et al. (1993, p. 204)). We have

$$\mu = \frac{\alpha}{\beta}, \qquad \sigma^2 = \frac{\alpha}{\beta^2}, \qquad \nu'_{3,\mu} = \frac{2\alpha}{\beta^3}, \qquad \nu_{4,\mu} = \frac{3\alpha(\alpha+2)}{\beta^4}.$$

4.2 Upper bounds for the total variation distance

In what follows, we give some upper bounds for

$$d_\tau := d_{TV}(\mathrm{NB}(q,\alpha), \Pi_\mu).$$

Using (6) with $\epsilon = 1$, we obtain

$$\mathrm{E}\Big(X \ln \frac{X}{\mu}\Big) \le \frac{\alpha}{\beta}\ln\Big(1+\frac{1}{\alpha}\Big).$$

Hence, by (4),

$$d_\tau \le \min\Big\{\frac{3\alpha}{2\beta e}\ln\Big(1+\frac{1}{\alpha}\Big), \frac{\alpha}{\beta^2}\Big\} \le \min\Big\{\frac{3}{2\beta e}, \frac{\alpha}{\beta^2}\Big\}. \qquad (30)$$

A sharper bound can be derived as follows: Observe that

$$\mathrm{E}\Big(X \ln \frac{X}{\mu}\Big) = \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty t^\alpha\, e^{-t\beta} \ln \frac{\beta\, t}{\alpha}\, dt = \frac{1}{\beta}\Big(1 + \alpha\Psi(\alpha) - \alpha \ln \alpha\Big),$$

where

$$\begin{aligned}\Psi(z) &= \frac{1}{\Gamma(z)}\frac{d}{dz}\Gamma(z) = \frac{1}{\Gamma(z)}\int_0^\infty t^{z-1}\, e^{-t}\, \ln t\, dt \\ &= \ln z - \frac{1}{2z} - 2\int_0^\infty \frac{t\,dt}{(t^2+z^2)(e^{2\pi t}-1)},\end{aligned}$$

for $z \in (0,\infty)$, is the digamma function. The latter equality can be found in Whittaker and Watson (1927, p. 251). Hence

$$\mathrm{E}\Big(X \ln \frac{X}{\mu}\Big) = \frac{1}{\beta}\Big(\frac{1}{2} - 2\alpha\int_0^\infty \frac{t\,dt}{(t^2+\alpha^2)(e^{2\pi t}-1)}\Big),$$

giving

$$\begin{aligned} d_\tau &\le \min\Big\{\frac{3}{2\beta e}\Big(\frac{1}{2} - 2\alpha\int_0^\infty \frac{t\,dt}{(t^2+\alpha^2)(e^{2\pi t}-1)}\Big), \frac{\alpha}{\beta^2}\Big\} \\ &\le \min\Big\{\frac{3}{4\beta e}, \frac{\alpha}{\beta^2}\Big\} = \sigma^2 \min\Big\{\frac{3}{4\mu e}, 1\Big\}, \qquad (31)\end{aligned}$$

which is better than (30).

4.3 Comparison with results in the literature

Let us have a look at the results given in the literature. For simplicity, we assume $\alpha \in \mathbf{N}$. Then

$$d_\tau \le \frac{1}{\beta}, \qquad \text{[Vervaat (1969, (1.8))]}, \qquad (32)$$

$$d_\tau \le \frac{1}{\beta\sqrt{2}}, \qquad \text{[Romanowska (1977, p. 129)]}, \qquad (33)$$

$$d_\tau \le \frac{\alpha(1-q)}{\beta^2}, \qquad \text{[Gerber (1984, Theorem 2(a))]}, \qquad (34)$$

$$d_\tau \le \frac{\alpha}{\beta^2}, \qquad \text{[Pfeifer (1987, (2.1))]}, \qquad (35)$$

$$d_\tau \le \frac{1-e^{-\alpha/\beta}}{\beta} \le \frac{1}{\beta}\min\Big\{\frac{\alpha}{\beta}, 1\Big\}, \qquad \text{[Barbour (1987, p. 758)]}. \qquad (36)$$

Vervaat (1969) proved his bound by using the Kullback–Leibler distance and Romanowska's (1977) note contains a refinement of Vervaat's method, leading to (33). This inequality can also be found in Matsunawa (1982, Corollary 3.2). Gerber (1984) proved his Theorem 2(a) by using the subadditivity property of d_{TV}. Pfeifer (1987) used a semigroup setting for his results. Barbour (1987) showed his bound (see (36)) by using the Stein–Chen method. Note that (36) can also be derived from the inequality due to Barbour et al. (1992) (see (2)). Observe that our inequalities (30) and (31) are better than (32), (33), and the second bound in (36). In particular, (30) has a better order, since $\alpha\ln(1+\alpha^{-1}) \to 0$ as $\alpha \to 0$. Gerber's inequality (see (34)) is better than Pfeifer's (1987) inequality (see (35)), which is, in turn, also contained in (31). Note that Lorz and Heinrich (1991, Example 3 and (30)) gave the bound $2\pi^{-1}\alpha\beta^{-2}$ for the Kolmogorov distance between $\mathrm{NB}(q,\alpha)$ and Π_μ. In this context, our Corollary 1 leads to the sharper bound $2^{-1}\alpha\beta^{-2}$.

4.4 Asymptotic relations for d_τ and proof of Theorem 1(b)

Now we treat asymptotic relations for d_τ. In view of (24), (25), and (26), we see that an absolute constant $c_2 \in (0,\infty)$ exists such that

$$\begin{aligned}\Big|\beta\sqrt{2\pi e}\,d_\tau - 1\Big| &\le c_2\Big(\frac{1}{\mu}+\frac{\mu}{\sigma^2}\Big(\frac{|\nu'_{3,\mu}|}{\mu^{3/2}}+\frac{\nu_{4,\mu}}{\mu^2}\Big)\Big)\\ &= c_2\Big(\frac{\beta}{\alpha}+\frac{2}{\sqrt{\alpha\beta}}+\frac{3}{\beta}+\frac{6}{\alpha\beta}\Big).\end{aligned} \tag{37}$$

Since the left-hand side of (37) is bounded by an absolute constant (see (31)), we obtain

$$\Big|\beta\sqrt{2\pi e}\,d_\tau - 1\Big| \le c_3\Big(\frac{\beta}{\alpha}+\frac{1}{\sqrt{\alpha\beta}}+\frac{1}{\beta}\Big), \tag{38}$$

where $c_3 \in (0,\infty)$ is also an absolute constant. Hence $d_\tau \sim 1/(\beta\sqrt{2\pi e})$ as $\beta\to\infty$ and $\beta/\alpha \to 0$. Pfeifer's (1987) result (see Remark 4 of the present paper) implies that

$$d_\tau = \frac{1}{\beta\sqrt{2\pi e}}\Big(1+O\Big(\frac{\beta}{\alpha}\Big)\Big), \tag{39}$$

if

$$\frac{\alpha^2}{\beta^3} = O(1), \qquad \frac{\alpha}{\beta}\to\infty. \tag{40}$$

From (38) we see that (39) is valid in the presence of the more general condition

$$\frac{\alpha}{\beta^3} = O(1), \qquad \frac{\alpha}{\beta^2} = O(1). \tag{41}$$

Indeed the implications (40) $\Rightarrow$ $(\alpha/\beta^3 \to 0,\ \alpha/\beta^2\to 0)$ $\Rightarrow$ (41) are valid and an easy example shows that (41) $\not\Rightarrow$ (40).

Proof of Theorem 1(b). We now discuss the sharpness of the constant $3/(2e)$ in (4). From (22) and (24) it follows that, if $\alpha = \beta$,

$$\Big|\frac{4\beta e}{3} d_\tau - 1\Big| \le \frac{4\beta e}{9}(2|\nu'_{3,\mu}| + \nu_{4,\mu}) = \frac{4e}{9\beta}\Big(7 + \frac{6}{\beta}\Big).$$

This yields $d_\tau \sim 3/(4\beta e)$ as $\alpha = \beta$ and $\beta \to \infty$. Therefore the constant $3/(4e)$ in (31) is sharp. Since (31) was derived from (4), the assertion follows. ■

5 Remaining proofs

First we give a proof of the lemma used in the introduction.

Proof of Lemma 1. Using the easy inequality $\ln x \le x - 1$, $(x \in (0,\infty))$, we obtain, for all $t, \epsilon \in (0,\infty)$,

$$\begin{aligned} \mathrm{E}\Big(X \ln \frac{X}{\mu}\Big) &= \frac{1}{\epsilon}\mathrm{E}\Big(X \ln \frac{(t\,X)^\epsilon}{\mu^\epsilon}\Big) - \mu\,\ln t \\ &\le \frac{1}{\epsilon}\mathrm{E}\Big(X\Big(\frac{(t\,X)^\epsilon}{\mu^\epsilon} - 1\Big)\Big) - \mu\,\ln t = \frac{t^\epsilon\,\mu_{1+\epsilon}}{\epsilon\,\mu^\epsilon} - \frac{\mu}{\epsilon} - \mu \ln t. \end{aligned}$$

If we consider the right–hand side as a function of t, it attains its minimum at $t = (\mu^{1+\epsilon}/\mu_{1+\epsilon})^{1/\epsilon}$, giving (6). The proof of (7) is clear. ■

For $k \in \mathbf{Z}_+$ and $s, t \in [0,\infty)$, let $H_{k,s,t}$ be the finite signed measure with counting density

$$h_{k,s,t}(m) = H_{k,s,t}(\{m\}) = \sum_{j=0}^{k} \frac{(t-s)^j}{j!}\,\Delta^j \pi(m,s), \qquad (m \in \mathbf{Z}_+).$$

Observe that, for all $t \in [0,\infty)$, $H_{0,s,t} = \Pi_s$ is the Poisson distribution with mean $s \in [0,\infty)$. Further, we have

$$g_{k,s}(m) = \mathrm{E}(h_{k,s,X}(m)), \qquad (k, m \in \mathbf{Z}_+,\ s \in [0,\infty)). \tag{42}$$

The proof of Theorem 2 requires some preparations in form of the following three lemmas. The next lemma forms the main tool in the argument of this paper. See Pfeifer (1985, Lemma 4.1) for a similar version in the context of operator semigroups. For the case $k = 2$, one can also consult Pfeifer (1987, proof of Theorem 1).

Lemma 2 *For $k, m \in \mathbf{Z}_+$, $s, t \in [0,\infty)$, we have*

$$\pi(m,t) - h_{k,s,t}(m) = \int_s^t \frac{(t-y)^k}{k!}\,\Delta^{k+1}\pi(m,y)\,dy.$$

In the following two lemmas, we are concerned with sharp norm estimates.

Lemma 3 *For $p \in \{1, \infty\}$, $n \in \mathbf{Z}_+$, and $t \in [0, \infty)$,*

$$\|\Delta^n \pi(\cdot, t)\|_p \leq \|\Delta^n \pi(\cdot, 0)\|_p = V_p^{(n)},$$

where

$$V_1^{(n)} = 2^n, \qquad V_\infty^{(n)} = \binom{n}{\lfloor n/2 \rfloor}, \qquad (n \in \mathbf{Z}_+).$$

The proof of the preceeding lemma is easy and therefore omitted. The next lemma introduces the constants $U_p^{(n)}$.

Lemma 4 *Let*

$$U_p^{(n)} = \sup_{t \in (0,\infty)} (t^{(n+1)/2 - 1/(2p)} \|\Delta^n \pi(\cdot, t)\|_p), \qquad (n \in \mathbf{Z}_+,\ p \in \{1, \infty\}).$$

Then the constants $U_p^{(n)}$ satisfy (13) *and* (14). *Moreover, for*

$$(p, n) \in \{(1,0),\ (1,1),\ (1,2),\ (\infty,0),\ (\infty,1),\ (\infty,2)\},$$

we have

$$\|\Delta^n \pi(\cdot, A_p^{(n)})\|_p = \frac{U_p^{(n)}}{(A_p^{(n)})^{(n+1)/2-1/(2p)}}, \tag{43}$$

where

$$A_1^{(0)} \in (0, \infty), \quad A_1^{(1)} = \frac{1}{2}, \quad A_1^{(2)} = 1, \quad A_\infty^{(0)} = \frac{1}{2}, \quad A_\infty^{(1)} = 1, \quad A_\infty^{(2)} = \frac{3}{2}.$$

Proof. Let us first mention that the finiteness of the $U_p^{(n)}$ is already known. Indeed, in the case $p = 1$ and $n \in \mathbf{N}$, Deheuvels and Pfeifer (1988) showed that $U_1^{(n)} \leq (2n/e)^{n/2}$, which is weaker than the first inequality in (14) for $n \geq 2$. The second inequality in (14), which corresponds to the case $p = \infty$, was proved by Shorgin (1977, proof of his Lemma 6). We now show the first inequality in (14). For this, we use (9), Cauchy's inequality, and the well-known fact that the Charlier polynomials are orthogonal with respect to the Poisson distribution (see, for example, Chihara (1978, (1.14), p. 4)):

$$\begin{aligned} \|\Delta^n \pi(\cdot, t)\|_1 &= \frac{1}{t^n} \sum_{m=0}^{\infty} \pi(m, t) |C_n(m, t)| \\ &\leq \frac{1}{t^n} \Big(\sum_{m=0}^{\infty} \pi(m, t) (C_n(m, t))^2 \Big)^{1/2} = \frac{\sqrt{n!}}{t^{n/2}}, \end{aligned}$$

for $n \in \mathbf{N}$ and $t \in (0, \infty)$. For (13) and (43), see Lemma 3 in Roos (2001) and observe that $2\, \|\pi(\cdot, t)\|_\infty = \|\Delta \pi(\cdot, t)\|_1$ for $t \in (0, \infty)$, (see Deheuvels and Pfeifer (1986b) or Roos (1999, (23) with $k = 1$)). ■

Proof of Theorem 2. Let $p \in \{1, \infty\}$, $s \in [0, \infty)$, $k \in \mathbf{Z}_+$, and $i \in \mathbf{Z}$ with

$k+i+1 \geq 0$. Using (10), (1), (42), the Lemmas 2–4, and Fubini's theorem in the case $p=1$, we obtain

$$\begin{aligned}
d_p^{(i)}(G,\, G_{k,s}) &= \Big\|\Delta^i(g - g_{k,s})\Big\|_p \\
&= \bigg\|\Delta^i\Big(\Big(\mathrm{E}[\pi(m,X) - h_{k,s,X}(m)]\Big)_{m\in\mathbf{Z}_+}\Big)\bigg\|_p \\
&= \bigg\|\Delta^i\Big(\Big(\mathrm{E}\Big[\int_s^X \frac{(X-y)^k}{k!}\Delta^{k+1}\pi(m,y)\,dy\Big]\Big)_{m\in\mathbf{Z}_+}\Big)\bigg\|_p \\
&= \bigg\|\Big(\mathrm{E}\Big[\int_s^X \frac{(X-y)^k}{k!}\Delta^{k+i+1}\pi(m,y)\,dy\Big]\Big)_{m\in\mathbf{Z}_+}\bigg\|_p \\
&\leq \mathrm{E}\bigg|\int_s^X \frac{|X-y|^k}{k!}\|\Delta^{k+i+1}\pi(\cdot,y)\|_p\,dy\bigg| \\
&\leq \mathrm{E}\bigg|\int_s^X \frac{|X-y|^k}{k!}\min\Big\{V_p^{(k+i+1)},\, \frac{U_p^{(k+i+1)}}{y^{(k+i+2)/2-1/(2p)}}\Big\}\,dy\bigg|.
\end{aligned}$$

This shows the validity of (12). Further, we see that (11) holds:

$$\begin{aligned}
d_p^{(i)}(G,\, G_{k,s}) &\leq \frac{V_p^{(k+i+1)}}{k!}\mathrm{E}\Big(\Big(\mathbf{1}_{(s,\infty)}(X) - \mathbf{1}_{[0,s)}(X)\Big)\int_s^X |X-y|^k\,dy\Big) \\
&= \frac{V_p^{(k+i+1)}}{k!}\mathrm{E}\Big(\mathbf{1}_{(s,\infty)}(X)\frac{(X-s)^{k+1}}{k+1} + \mathbf{1}_{[0,s)}(X)\frac{(s-X)^{k+1}}{k+1}\Big) \\
&= \frac{V_p^{(k+i+1)}}{(k+1)!}\nu_{k+1,s},
\end{aligned}$$

where, for a set B, $\mathbf{1}_B(t) = 1$ if $t \in B$ and $\mathbf{1}_B(t) = 0$ otherwise. ∎

Proof of Corollary 1 and Theorem 1(a). First of all, we show (16) directly: We have $\Pi_\mu = G_{1,\mu}$ and therefore, as in the proof of Theorem 2,

$$\begin{aligned}
d_1^{(-2)}(G,\, \Pi_\mu) &= d_1^{(-2)}(G,\, G_{1,\mu}) \\
&= \bigg\|\Big(\mathrm{E}\Big[\int_\mu^X (X-y)\pi(m,y)\,dy\Big]\Big)_{m\in\mathbf{Z}_+}\bigg\|_1.
\end{aligned}$$

Since $\int_\mu^X (X-y)\pi(m,y)\,dy \geq 0$ for all $m \in \mathbf{Z}_+$, we obtain

$$\begin{aligned}
d_1^{(-2)}(G,\, \Pi_\mu) &= \sum_{m=0}^\infty \mathrm{E}\Big(\int_\mu^X (X-y)\pi(m,y)\,dy\Big) \\
&= \mathrm{E}\Big(\int_\mu^X (X-y)\,dy\Big) = \frac{\sigma^2}{2}.
\end{aligned}$$

In what follows, we only prove (4). The other inequalities are shown in the same way. We use Theorem 2 with $k=1$ and derive:

$$\begin{aligned}
d_{TV}(G,\, \Pi_\mu) &= d_{TV}(G,\, G_{1,\mu}) \\
&\leq \frac{U_1^{(2)}}{2}\mathrm{E}\Big(\int_\mu^X \frac{X-y}{y}\,dy\Big) = \frac{3}{2e}\mathrm{E}\Big(X\ln\frac{X}{\mu}\Big).
\end{aligned}$$

Together with the first inequality in (11), this leads to (4). ■

For the proof of Theorem 3, we need the following lemma.

Lemma 5 *An absolute constant $c_4 \in (0, \infty)$ exists such that, for all $t \in (0, \infty)$,*

$$\left| \|\Delta^2 \pi(\cdot, t)\|_1 - \frac{4}{t\sqrt{2\pi e}} \right| \leq \frac{c_4}{t^2}. \tag{44}$$

Proof. In Roos (2000, (44)) it was proved that an absolute constant $c_4 \in (0, \infty)$ exists such that, for all $q \in (0, 1)$ and $n \in \{2, 3, 4, \ldots\}$,

$$\left| \|\Delta^2 b(\cdot, n-2, q)\|_1 - \frac{4}{nq(1-q)\sqrt{2\pi e}} \right| \leq \frac{c_4}{(nq(1-q))^2},$$

where $b(\cdot, n-2, q) \in \mathbf{R}^{\mathbf{Z}_+}$ with $b(m, n-2, q) = \binom{n-2}{m} q^m (1-q)^{n-2-m}$ for $m \in \mathbf{Z}_+$. By the triangle inequality, we obtain, for $t \in (0, \infty)$ and all $n \in \{3, 4, 5, \ldots\}$, $q \in (0, 1)$ with $(n-2)q = t$,

$$\begin{aligned} \left| \|\Delta^2 \pi(\cdot, t)\|_1 - \frac{4}{t\sqrt{2\pi e}} \right| &\leq \|\Delta^2(\pi(\cdot, t) - b(\cdot, n-2, q))\|_1 \\ &+ \left| \|\Delta^2 b(\cdot, n-2, q)\|_1 - \frac{4}{nq(1-q)\sqrt{2\pi e}} \right| + \frac{4}{\sqrt{2\pi e}} \left| \frac{1}{nq(1-q)} - \frac{1}{t} \right|. \end{aligned}$$

Since the first and the third term on the right-hand side tend to zero as $n \to \infty$ and $q \to 0$, the assertion follows. ■

Note that Proposition 4 in Roos (1999) leads to an asymptotic relation, which is weaker than (44).

Proof of Theorem 3. To prove (a), we use (21) with $p = 1$, $s = \mu$, $k = 1$, and $i = 0$, Theorem 2, Lemma 4, and the equality

$$\|\Delta^2 \pi(\cdot, t)\|_1 = \frac{2}{t} \Big(L_{b_+(t)}(t) - L_{b_-(t)}(t) \Big), \qquad (t \in (0, \infty))$$

(see Deheuvels and Pfeifer (1986a)). In particular, using (12), we obtain one part of (23):

$$\begin{aligned} d_{TV}(G,\, G_{2,\mu}) &\leq \frac{U_1^{(3)}}{4} \mathrm{E} \left| \int_\mu^X \frac{(X-y)^2}{y^{3/2}}\, dy \right| \\ &= \frac{U_1^{(3)}}{6\sqrt{\mu}} \mathrm{E}\Big(|\sqrt{X} - \sqrt{\mu}|^3 (3\sqrt{X} + \sqrt{\mu}) \Big) \\ &\leq \frac{U_1^{(3)}}{2\sqrt{\mu}} \mathrm{E}\Big(\frac{|X-\mu|^3}{(\sqrt{X} + \sqrt{\mu})^2} \Big). \end{aligned}$$

The other part of (23) follows from (11). Now we prove (24). Some preparations are needed. By (12) and (15), we have

$$d_{TV}(G,\, G_{3,\mu}) \leq \frac{U_1^{(4)}}{12} \mathrm{E}\Big((X-\mu)^4 \int_0^1 \frac{(1-y)^3\, dy}{(yX + (1-y)\mu)^2} \Big),$$

where

$$\int_0^1 \frac{(1-y)^3\,dy}{(yX+(1-y)\mu)^2} \le \int_0^1 \frac{(1-y)\,dy}{(yX+\mu)^2} = \frac{1}{\mu X} + \frac{1}{X^2}\ln\frac{\mu}{X+\mu} \le \frac{1}{\mu(X+2\mu)}.$$

For the latter estimate, we used the inequality $\ln x \le 2(x-1)/(x+1)$, ($x \in (0,1]$), see Mitrinović (1970, 3.6.18). Together with (11), we now obtain

$$d_{TV}(G,\, G_{3,\mu}) \le \min\left\{\frac{U_1^{(4)}}{12\,\mu}\,\mathrm{E}\Big(\frac{(X-\mu)^4}{X+2\mu}\Big),\ \frac{1}{3}\,\nu_{4,\mu}\right\}. \tag{45}$$

Using (21) with $p=1$, $s=\mu$, $k=2$, and $i=0$, (45), Lemmas 3 and 4, we are led to (24):

$$\begin{aligned}
d_{TV}(G,\, G_{2,\mu}) &\le \frac{|\nu'_{3,\mu}|}{12}\,\|\Delta^3\pi(\cdot,\mu)\|_1 + d_{TV}(G,\, G_{3,\mu}) \\
&\le \frac{|\nu'_{3,\mu}|}{12}\,\min\left\{\frac{U_1^{(3)}}{\mu^{3/2}},\, 8\right\} + \min\left\{\frac{U_1^{(4)}}{12\,\mu}\,\mathrm{E}\Big(\frac{(X-\mu)^4}{X+2\mu}\Big),\ \frac{1}{3}\,\nu_{4,\mu}\right\} \\
&\le \min\left\{\frac{U_1^{(3)}\,|\nu'_{3,\mu}|}{12\,\mu^{3/2}} + \frac{U_1^{(4)}}{12\,\mu}\mathrm{E}\Big(\frac{(X-\mu)^4}{X+2\mu}\Big),\ \frac{1}{3}\,(2|\nu'_{3,\mu}| + \nu_{4,\mu})\right\}.
\end{aligned}$$

Assertion (b) follows from (21) and (44):

$$\begin{aligned}
\left|d_{TV}(G,\,\Pi_\mu) - \frac{\sigma^2}{\mu\sqrt{2\pi e}}\right| &\le \left|d_{TV}(G,\,\Pi_\mu) - \frac{\sigma^2}{4}\,\|\Delta^2\pi(\cdot,\mu)\|_1\right| \\
&\quad + \frac{\sigma^2}{4}\left|\|\Delta^2\pi(\cdot,\mu)\|_1 - \frac{4}{\mu\sqrt{2\pi e}}\right| \\
&\le d_{TV}(G,\, G_{2,\mu}) + \frac{c_4\,\sigma^2}{4\,\mu^2}.
\end{aligned}$$

The theorem is proved. ■

References

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

[2] Barbour, A.D., Holst, L., Janson, S. (1992). Poisson Approximation. Clarendon Press, Oxford.

[3] Chihara, T.S. (1978). An Introduction to Orthogonal Polynomials. Gordon and Breach, New York.

[4] Daley, D.J., Vere-Jones, D. (1988). An Introduction to the Theory of Point Processes. Springer Verlag, New York.

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

[6] Deheuvels, P., Pfeifer, D. (1986b). *Operator semigroups and Poisson convergence in selected metrics.* Semigroup Forum, **34**, 203–224. Errata: Semigroup Forum, **35**, 251, (1987).

[7] Deheuvels, P., Pfeifer, D. (1988). *On a relationship between Uspensky's theorem and Poisson approximation.* Ann. Inst. Statist. Math., **40**, 671–681.

[8] DOUGLAS, J.B. (1980). Analysis with Standard Contagious Distributions. International Co-operative Publishing House, Fairland, Maryland.

[9] FELLER, W. (1943). *On a general class of contagious distributions.* Ann. Math. Statist., **14**, 389–400.

[10] FREEDMAN, D. (1974). *The Poisson approximation for dependent events.* Ann. Probab., **2**, 256–269.

[11] GERBER, H.U. (1984). *Error bounds for the compound Poisson approximation.* Insurance Math. Econom., **3**, 191–194.

[12] GRANDELL, J. (1997). Mixed Poisson Processes. Chapman & Hall, London.

[13] HAIGHT, F.A. (1967). Handbook of the Poisson Distribution. John Wiley, New York.

[14] JOHNSON, N.L., KOTZ, S., KEMP, A.W. (1993). Univariate Discrete Distributions. Second Edition. John Wiley, New York.

[15] LORZ, U., HEINRICH, L. (1991). *Normal and Poisson approximation of infinitely divisible distribution functions.* Statistics, **22**, 627–649.

[16] MATSUNAWA, T. (1982). *Some strong epsilon-equivalence of random variables.* Ann. Inst. Stat. Math., **34**, 209–224.

[17] MITRINOVIĆ, D.S. (1970). Analytic Inequalities. Springer Verlag, Berlin.

[18] PFEIFER, D. (1985). *Approximation-theoretic aspects of probabilistic representations for operator semigroups.* J. Approximation Theory, **43**, 271–296.

[19] PFEIFER, D. (1987). *On the distance between mixed Poisson and Poisson distributions.* Statistics & Decisions, **5**, 367–379.

[20] RACHEV, S.T. (1991). Probability Metrics and Stability of Stochastic Models. John Wiley, Chichester.

[21] 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.

[22] ROOS, B. (1999). *Asymptotics and sharp bounds in the Poisson approximation to the Poisson binomial distribution.* Bernoulli, **5**, 1021–1034.

[23] ROOS, B. (2000). *Binomial approximation to the Poisson binomial distribution: The Krawtchouk expansion.* Theory Probab. Appl., **45**, 258–272.

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

[25] SHORGIN, S.Y. (1977). *Approximation of a generalized binomial distribution.* Theory Probab. Appl., **22**, 846–850.

[26] 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.

[27] WHITTAKER, E.T., WATSON, G.N. (1927). A Course of Modern Analysis. 4th Ed., Cambridge University Press, London.

[28] YANNAROS, N. (1991). *Poisson approximation for random sums of Bernoulli random variables.* Statist. Probab. Lett., **11**, 161–165.