Statistics & Decisions 5, 367–379 (1987)

ON THE DISTANCE BETWEEN MIXED POISSON AND POISSON DISTRIBUTIONS

D. Pfeifer*)

Received: Revised version: 15.12.1986

Summary. Estimations and asymptotic expansions for several distances between mixed Poisson and Poisson distributions are given, such as the total variation distance, the Kolmogorov distance and a specific Wasserstein distance (Fortet-Mourier distance). As an example, we generalize and improve results of Vervaat [9] on the total variation distance between negative binomial and Poisson distributions. The main tool is an appropriate application of operator semigroups and their probabilistic representation theory.

1. The Semigroup Setting of Poisson Approximation

Let $\mathfrak{M}$ denote the set of all probability measures over $\mathbb{Z}^+$, the non-negative integers, and let P^S, $P^T \in \mathfrak{M}$ denote the distribution of discrete random variables S and T. On $\mathfrak{M}$, metrics d_i ($i = 1,2,3$) will be considered given by

---

*) Research supported in part by US AFOSR Contract No. F49620 82 C 0009

AMS subject classifications: Primary 62 E 20, 60 F 05, Secondary 47 D 05

Keywords: Poisson approximation, negative Poisson binomial distributions, operator semigroups, total variation distance, Kolmogorov distance, Wasserstein distance

(1.1) $$d_1(P^S,P^T) = \sup_{M\subseteq\mathbb{Z}^+} |P(S\in M) - P(T\in M)| = \frac{1}{2}\sum_{k=0}^{\infty}|P(S=k)-P(T=k)|$$

(total variation distance)

(1.2) $$d_2(P^S,P^T) = \sup_{k\in\mathbb{Z}^+} |F_S(k) - F_T(k)|,$$

where F_S, F_T denotes the c.d.f. of S and T (Kolmogorov distance)

(1.3) $$d_3(P^S,P^T) = \inf_Q E(|S-T|) = \sum_{k=0}^{\infty}|F_S(k) - F_T(k)|,$$

where Q ranges over all joint distributions of (S,T) with the given marginals (Wasserstein or Fortet-Mourier distance); see Vallender [8].

These metrics can in a natural way be imbedded in a more general Banach space setting as follows.

Let $\mathfrak{X}$ denote the Banach space ℓ^1 of all real-valued absolutely summable sequences $f = (f(0),f(1),\dots)$ or the Banach space ℓ^∞ of all absolutely bounded sequences, resp. A measure $\mu\in\mathfrak{M}$ will be identified with the element $(\mu(\{0\}),\mu(\{1\}),\dots)\in\ell^1$. For $f\in\ell^1$, $g\in\ell^\infty$ the convolution $f*g$ is defined by

(1.4) $$f*g(n) = \sum_{k=0}^{n} f(k)g(n-k),\ n\in\mathbb{Z}^+ .$$

Then again, $f*g\in\mathfrak{X}$, and we have

(1.5) $$\|f*g\|_{\mathfrak{X}} \le \|f\|_{\ell^1}\|g\|_{\mathfrak{X}} ,$$

where $\|\cdot\|_{\mathfrak{X}}$ denotes the corresponding norm on $\mathfrak{X}$. Any measure $\mu\in\mathfrak{M}$ can now be considered as a bounded linear operator on $\mathfrak{X}$ via

(1.6) $$\mu g = \mu * g,\ g\in\mathfrak{X}.$$

In fact, μ is a positive contraction on $\mathfrak{X}$ with

(1.7) $$\|\mu\| = \sum_{k=0}^{\infty}\mu(\{k\}) = 1.$$

The metrics d_i, $i=1,2,3$ can then be given the following equivalent form

(1.8) $d_1(\mu,\nu) = \frac{1}{2}\,||(\mu-\nu)g\,||_{\ell^1},\ g = (1,0,0,\dots)$

(1.9) $d_2(\mu,\nu) = ||(\mu-\nu)h\,||_{\ell^\infty},\ h = (1,1,1,\dots)$

(1.10) $d_3(\mu,\nu) = \lim_{n\to\infty} ||(\mu-\nu)h_n\,||_{\ell^1},\ h_n = (1,\dots,1,0,0,\dots)$

(containing n times 1 and 0 otherwise), for all $\mu,\nu\in\mathfrak{M}$. If specifically μ is a Poisson distribution with mean $\xi\geq 0$, then the semigroup representation

(1.11) $$\mu = e^{A\xi} = \sum_{k=0}^{\infty} \frac{\xi^k}{k!} A^k$$

holds where the infinitesimal generator A is defined by

(1.12) $$Ag(n) = \begin{cases} g(n-1)-g(n), & n\geq 1 \\ -g(0), & n=0\ . \end{cases}$$

Obviously, A is a bounded linear operator on $\mathfrak{X}$ with $||A\,|| = 2$ in all cases. Now if μ is a mixed Poisson distribution with mixing r.v. X i.e. $\mu(\{k\}) = \int_0^\infty e^{-k\xi}\frac{\xi^k}{k!}P^X(d\xi)$ for $k\geq o$, or in the above operator setting

(1.13) $$\mu = E(e^{AX}) = \int_0^\infty e^{A\xi}P^X(d\xi)$$

which by the uniform continuity of the semigroup exists as a Pettis integral in $\mathfrak{L}(\mathfrak{X})$, the Banach algebra of all bounded endomorphisms on $\mathfrak{X}$, where the composition of mappings plays the role of the (Banach) product. For the estimation of the distances d_i it is therefore necessary to analyze the norm distances

(1.14) $$||E(e^{AX})g - e^{A\xi}g\,||_{\mathfrak{X}}$$

for elements $g\in\mathfrak{X}$. This can be done in generality by means of the probabilistic representation theory of operator semigroups as developed in Pfeifer [6], [7].

Actually, relationships between convolutions and semigroups have been investigated earlier (see LeCam [5] and Feller [4]), even for more general Banach spaces. However, for our purposes, we shall need estimations independent from the convolution structure, such as the following (main) result.

THEOREM 1. Let $\mathfrak{X}$ be a Banach space (which may be arbitrary) and A the (bounded) generator of a contraction semigroup on $\mathfrak{X}$. Let further X be a non-negative real-valued random variable. Then for all $g \in \mathfrak{X}$, the following estimations hold true.

(1.15) $$\|E(e^{AX})g - e^{A\xi}g\|_{\mathfrak{X}} \leq \frac{1}{2} E(X-\xi)^2 \|A^2 g\|_{\mathfrak{X}} \text{ if } X \in L^2(P) \text{ and } E(X) = \xi;$$

(1.16) $$\left| \|E(e^{AX})g - e^{A\xi}g\|_{\mathfrak{X}} - \|E(X-\xi)e^{A\xi}Ag + \frac{1}{2}E(X-\xi)^2 e^{A\xi}A^2 g\|_{\mathfrak{X}}\right| \leq \frac{1}{6}\|A^3 g\|_{\mathfrak{X}} E(|X-\xi|^3), \text{ if } X \in L^3(P);$$

(1.17) $$\left| \|E(e^{AX})g - e^{A\xi}g\|_{\mathfrak{X}} - \|E(X-\xi)e^{A\xi}Ag + \frac{1}{2}E(X-\xi)^2 e^{A\xi}A^2 g\|_{\mathfrak{X}}\right| \leq \frac{1}{6}\|e^{A\xi/2}A^3 g\|_{\mathfrak{X}} E(|X-\xi|^3) + \frac{1}{3\xi}\|A^3 g\|_{\mathfrak{X}} E(X-\xi)^4, \text{ if } X \in L^4(P).$$

Here throughout, $\xi > 0$.

Proof. Theorem 1 is basically an extension of Theorem 4.1 in Pfeifer [7], the proof hence follows the same lines. We shall give here a detailed proof of (1.17) only. Since

$$E(e^{AX})g - e^{A\xi}g = E[(X-\xi)e^{A\xi}Ag + \frac{1}{2}(X-\xi)^2 e^{A\xi}A^2 g + \int_\xi^X \frac{(X-u)^2}{2} e^{Au}A^3 g\, du],$$

the remainder term in (1.17) is obtained by a suitable estimation of the expectation of the latter integral. If $I(\cdot)$ denotes the indicator r.v. of the event specified, we have

$$\begin{aligned} E[\int_\xi^X \frac{(X-u)^2}{2} e^{Au}A^3 g\, du] &= E[I(X>\xi) \int_\xi^X \frac{(X-u)^2}{2} e^{A(u-\frac{\xi}{2})} e^{A\xi/2}A^3 g\, du \\ &- I(\xi/2 < X \leq \xi) \int_X^\xi \frac{(X-u)^2}{2} e^{A(u-\frac{\xi}{2})} e^{A\xi/2}A^3 g\, du \\ &- I(X \leq \xi/2) \int_X^\xi \frac{(X-u)^2}{2} e^{Au}A^3 g\, du], \end{aligned}$$

giving

$$\begin{aligned} \|E[\int_\xi^X \frac{(X-u)^2}{2} e^{Au}A^3 g\, du]\|_{\mathfrak{X}} &\leq E[I(X>\xi) \int_\xi^X \frac{(X-u)^2}{2} \|e^{A\xi/2}A^3 g\|_{\mathfrak{X}} du \\ &+ I(\xi/2 < X \leq \xi) \int_X^\xi \frac{(X-u)^2}{2} \|e^{A\xi/2}A^3 g\|_{\mathfrak{X}} du + I(X \leq \xi/2)\|A^3 g\|_{\mathfrak{X}} \frac{|X-\xi|^3}{6}] \\ &\leq \frac{1}{6} E(|X-\xi|^3)\|e^{A\xi/2}A^3 g\|_{\mathfrak{X}} + \frac{1}{6}\|A^3 g\|_{\mathfrak{X}} E(I(X \leq \xi/2)|X-\xi|^3). \end{aligned}$$

Now by Hölder's and Markov's inequality,

$$E[I(X \leq \xi/2)|X-\xi|^3] \leq (P(X \leq \xi/2))^{1/4}(E(X-\xi)^4)^{3/4} \leq \frac{2}{\xi} E(X-\xi)^4,$$

which finishes the proof.

Clearly, the right hand side of (1.17) can further be improved by e.g. exponential bounds for $(P(X \leq \xi/2))^{1/4}$, for instance if the moment generating function of X exists.

Theorem 1 is the key for suitable estimations for $d_i(\mu,\nu)$ in relations (1.8) to (1.10). This will be worked out in more detail in the subsequent chapters.

2. Approximation of Mixed Poisson Distributions - The Unbiased Case

Here we shall assume that for the mixing r.v. X we have $E(X) = \xi$ (which will be referred to as the unbiased case, since then the variance of X,σ^2, plays the central role in (1.15) to (1.17)). The following estimation for the distances $d_i(\mu,\nu)$ can immediately be derived from Theorem 1.

COROLLARY 1. Let ν be a mixed Poisson distribution with mixing r.v. X such that $E(X)=\xi$, and μ a Poisson distribution with mean $\xi>0$. Then

(2.1) $\quad d_i(\mu,\nu) \leq K_i\sigma^2$, if $X \in L^2(P)$,

where $K_1 = K_3 = 1$, $K_2 = \frac{1}{2}$,

(2.2) $\quad |d_i(\mu,\nu) - C_i \frac{\sigma^2}{2} \|e^{A\xi}A^2 g_i\|_{\mathfrak{X}_i}| \leq \frac{2}{3} K_i E(|X-\xi|^3)$, if $X \in L^3(P)$, where $C_1 = \frac{1}{2}$, $C_2 = C_3 = 1$, $\mathfrak{X}_1 = \mathfrak{X}_3 = \ell^1$, $\mathfrak{X}_2 = \ell^\infty$, $g_1 = (1,0,0,\ldots)$, $g_2 = g_3 = (1,1,1,\ldots)$;

if additionally also $X \in L^4(P)$, then the right hand side of (2.2) can also be replaced by

(2.3) $\quad C_i/6 \|e^{A\xi/2}A^3 g_i\|_{\mathfrak{X}_i} E(|X-\xi|^3) + \frac{4}{3} K_i/\xi E(X-\xi)^4$.

Note that here for d_3 no longer limit relations have to be considered, since now $A^2 g_3 = (1,-1,0,0,\ldots) \in \ell^1$, $A^3 g_3 = (-1,2,-1,0,0,\ldots) \in \ell^1$.

Corollary 1 allows for sharp upper and lower bounds for the distances d_i in the low range for ξ (relation (2.1)), in the moderate range (relation (2.2)) and the high range (relation (2.3)), in the latter cases provided that $E(|X-\xi|^3)$ and $E(X-\xi)^4$ are small enough in comparison with σ^2, resp., since the norm term in (2.3) can roughly be estimated by

$$(2.4) \qquad \|e^{A\xi/2}A^3 g_i\|_{\mathfrak{X}_i} \le \|A\| \; \|e^{A\xi/2}A^2 g_i\|_{\mathfrak{X}_i} = 2\,\|e^{A\xi/2}A^2 g_i\|_{\mathfrak{X}_i} .$$

An exact evaluation of the norm terms in Corollary 1 is given below.

THEOREM 2. With the notation of Corollary 1, we have for all $t > 0$

$$(2.5) \qquad \|e^{At}A^2 g_1\|_{\mathfrak{X}_1} = 2\left\{\frac{t^{a-1}(a-t)}{a!} + \frac{t^{b-1}(t-b)}{b!}\right\} e^{-t},$$

$$(2.6) \qquad \|e^{At}A^2 g_2\|_{\mathfrak{X}_2} = \max\left\{\frac{t^{a-1}(a-t)}{a!}, \frac{t^{b-1}(t-b)}{b!}\right\} e^{-t},$$

$$(2.7) \qquad \|e^{At}A^2 g_3\|_{\mathfrak{X}_3} = \frac{t^{[t]}}{[t]!} e^{-t},$$

where $[\cdot]$ denotes the integer part of the real number specified, and

$$(2.8) \qquad a = [t + \tfrac{1}{2} + \sqrt{t + 1/4}], \; b = [t + \tfrac{1}{2} - \sqrt{t+1/4}] .$$

Further,

$$(2.9) \qquad \|e^{At}A^3 g_2\|_{\mathfrak{X}_2} \le \frac{1}{t}\,\frac{t^{[t]}}{[t]!} e^{-t} + \frac{1}{t\sqrt{t}}$$

$$(2.10) \qquad \|e^{At}A^3 g_3\|_{\mathfrak{X}_3} = \|e^{At}A^2 g_1\|_{\mathfrak{X}_1} .$$

Proof. Relation (2.5) was proved in Deheuvels and Pfeifer [2], while (2.6) and (2.7) were proved in Deheuvels and Pfeifer [3]. For the proof of (2.9) and (2.10) we only have to observe that $A^3 g_2 = A^3 g_3 = -A^2 g_1$, from which (2.10) is immediately obvious. Also,

$$\|e^{At}A^2 g_1\|_{\ell^\infty} = \sup_{n\geq 0} e^{-t}\frac{t^{n-2}}{n!}\,|t^2 - 2nt + n(n-1)|$$

$$(2.11)\qquad = \sup_{n\geq 0} e^{-t}\frac{t^{n-2}}{n!}\,|(n-t)^2 - n|$$

$$\leq \frac{1}{t}\sup_{n\geq 1} e^{-t}\frac{t^{n-1}}{(n-1)!} + \frac{1}{t^2}\sup_{n\geq 0} e^{-t}\frac{t^n}{n!}(n-t)^2,$$

from which (2.9) follows by Stirling's formula.

It should be pointed out that the expressions on the right hand side of (2.5) and (2.6) can also be represented as

$$(2.12)\qquad \frac{t^{a-1}(a-t)}{a!} = \frac{t^{k-1}(k-t)}{k!} \text{ whenever } k-\sqrt{k} \leq t < k+1-\sqrt{k+1},\ k \geq 1,$$

$$(2.13)\qquad \frac{t^{n-1}(t-b)}{b!} = \frac{t^{k-1}(t-k)}{k!} \text{ whenever } k+\sqrt{k} \leq t < k+1+\sqrt{k+1},\ k \geq 0.$$

In fact, part of these developments (however for the Poisson binomial setting) have been given by Barbour and Hall [1], p. 477, who suggested that the quantities above would be "not in general very neatly expressible". Also, by Stirling's formula, we have

$$(2.14)\qquad \lim_{t\to\infty} \frac{t}{4}\|e^{At}A^2 g_1\|_{\mathfrak{X}_1} = (2\pi e)^{-1/2} \sim .242,$$

a fact that has also been observed by Barbour and Hall [1], apparently without recognizing the constants π and e being involved here.

Actually, by use of Stirling's formula, we can give the following bounds on (2.12) and (2.13).

<u>LEMMA.</u> We have

$$(2.15)\qquad \frac{1}{t}(2\pi e)^{-1/2}\exp(-3/(2\sqrt{t})) \leq \frac{t^{a-1}(a-t)}{a!}e^{-t} \leq \frac{1}{t}(2\pi e)^{-1/2},\ t \geq 3-\sqrt{3},$$

$$(2.16)\qquad \frac{1}{t}(2\pi e)^{-1/2}\exp(-2/\sqrt{t}) \leq \frac{t^{b-1}(b-t)}{b!}e^{-t} \leq \frac{1}{t}(2\pi e)^{-1/2}\exp(2/(2\sqrt{t})),\quad t \geq 3+\sqrt{3},$$

$$(2.17)\qquad (2\pi t)^{-1/2}\exp(-3/(2t)) \leq \frac{t^{[t]}}{[t]!}e^{-t} \leq (2\pi t)^{-1/2}\sqrt{1+2/t},\ t \geq 1.$$

<u>Proof.</u> Let $H(t) = \frac{t^k(k-t)}{k!}$ for $k-\sqrt{k} \leq t < k+1-\sqrt{k+1}$, $k \geq 3$. In that

range, $H(t)$ is maximal for $t = k + \frac{1}{2} - \sqrt{k+1/4}$ and minimal for $t = k - \sqrt{k}$. Similarly, if $J(t) = \frac{t^k(t-k)}{k!} e^{-t}$ for $k + \sqrt{k} \leq t < k + 1 + \sqrt{k+1}$, $k \geq 3$, then $J(t)$ is maximal in that range for $t = k + 1/2 + \sqrt{k+1/4}$, and minimal for $t = k + 1 + \sqrt{k+1}$. By the use of a little analysis and Stirling's formula, (2.15) to (2.17) are now easily proved.

By means of Corollary 1, Theorem 2 and the above Lemma we have at once a large number of upper and lower bounds for the distances $d_i(\mu,\nu)$ available. Note that also for $t > 0$

$$\text{(2.18)}\quad \begin{aligned} \|e^{At}A^3 g_1\|_{\mathfrak{X}_1} &= e^{-t} \sum_{k=0}^{\infty} \frac{t^{k-3}}{k!} \left| t^3 - 3kt^2 + 3k(k-1)t - k(k-1)(k-2) \right| \\ &\leq \frac{1}{t^3} E(|T-t|^3) + \frac{3}{t^2} E(|T-t|) + \frac{1}{t^3} E(T) \end{aligned}$$

where T is a Poisson random variable with mean t, giving

$$\text{(2.19)}\quad \|e^{At}A^3 g_1\|_{\mathfrak{X}_1} \leq \frac{\sqrt{1+3t}}{2^t} + \frac{3}{t\sqrt{t}} + \frac{1}{t^2} = 0\left(\frac{1}{t\sqrt{t}}\right)$$

for $t \to \infty$ which is the same rate as in the second order estimations in Barbour and Hall [1] for the Poisson binomial setting.

An interesting asymptotic result is given by

COROLLARY 2. Suppose that the distribution of X depends on the mean ξ and the variance σ^2 in such a way that
$E(X-\xi)^4 = 0\left(\frac{\sigma^2}{\xi}\right)$ for $\xi \to \infty$. Then

$$\text{(2.20)}\quad d_1(\mu,\nu) = \frac{\sigma^2}{\xi}(2\pi e)^{-1/2} + 0\left(\frac{\sigma^2}{\xi^2}\right),$$

$$\text{(2.21)}\quad d_2(\mu,\nu) = \frac{\sigma^2}{2\xi}(2\pi e)^{-1/2} + 0\left(\frac{\sigma^2}{\xi^2}\right),$$

$$\text{(2.22)}\quad d_3(\mu,\nu) = \frac{\sigma^2}{2}(2\pi\xi)^{-1/2} + 0\left(\frac{\sigma^2}{\xi\sqrt{\xi}}\right).$$

Note that in Corollary 2, we necessarily have $\sigma^2 \to 0$ since $\sigma^4 \leq E(X-\xi)^4$, hence $\sigma^2 = 0(\frac{1}{\xi})$.

An important example of mixed Poisson distributions are the negative

Poisson binomial distributions $\nu = \mathop{*}\limits_{k=1}^{n} \nu_k$ where each ν_k is a geometric distribution with $\nu_k(\{0\}) = p_k$ where $0 < p_k = 1 - q_k < 1$. In this case, we may choose $X = \sum_{k=1}^{n} X_k$ where $X_1, \ldots, X_n$ are independent exponentially distributed r.v.'s with means q_k/p_k each since

$$(2.23) \qquad \nu = E(e^{A \sum_{k=1}^{n} X_k}) = E(\prod_{k=1}^{n} e^{AX_k}) = \prod_{k=1}^{n} E(e^{AX_k}) = \mathop{*}\limits_{k=1}^{n} \nu_k$$

(cf. also Feller [4], p. 573). Also,

$$(2.24) \qquad \nu_k = E(e^{AX_k}) = \frac{q_k}{p_k} R(\frac{q_k}{p_k}) = (I - \frac{p_k}{q_k} A)^{-1}$$

where R denotes the resolvent of the semigroup and I the identity operator (see Deheuvels and Pfeifer [2], or Pfeifer [7]), which gives a justification for the name "negative Poisson binomial distribution" for ν. Here,

$$(2.25) \qquad \xi = \sum_{k=1}^{n} \frac{q_k}{p_k}, \quad \sigma^2 = \sum_{k=1}^{n} (\frac{q_k}{p_k})^2,$$

$$(2.26) \qquad E(X - \xi)^4 = 6 \sum_{k=1}^{n} (\frac{q_k}{p_k})^4 + 3\sigma^4 \leqq 9\sigma^4,$$

which follows easily by the relation

$$(2.27) \qquad E(X - \xi)^4 = G^{(4)}(0) + 3\sigma^4,$$

where $G(s) = \log(E(e^{sX}))$, $s \geqq 0$.

Now if the q_k depend on n in such a way that $\max(q_1, \ldots, q_n) \to 0$, $\sum_{k=1}^{n} q_k \to \infty$, $(\sum_{k=1}^{n} q_k)(\sum_{k=1}^{n} q_k^2) = O(1)$ for $n \to \infty$, then by Corollary 2,

$$(2.28) \qquad d_1(\mu, \nu) = \frac{\sum_{k=1}^{n} (\frac{q_k}{p_k})^2}{\sum_{k=1}^{n} \frac{q_k}{p_k}} (2\pi e)^{-1/2} + O((\sum_{k=1}^{n} q_k^2)(\sum_{k=1}^{n} q_k)^{-2}),$$

$$(2.29)\qquad d_2(\mu,\nu) = \frac{1}{2}\frac{\sum_{k=1}^{n}\left(\frac{q_k}{p_k}\right)^2}{\sum_{k=1}^{n}\frac{q_k}{p_k}}(2\pi e)^{-1/2} + 0\left(\left(\sum_{k=1}^{n} q_k^2\right)\left(\sum_{k=1}^{n} q_k\right)^{-2}\right),$$

$$(2.30)\qquad d_3(\mu,\nu) = \frac{1}{2}\frac{\sum_{k=1}^{n}\left(\frac{q_k}{p_k}\right)^2}{\sqrt{\sum_{k=1}^{n}\frac{q_k}{p_k}}}(2\pi)^{-1/2} + 0\left(\left(\sum_{k=1}^{n} q_k^2\right)\left(\sum_{k=1}^{n} q_k\right)^{-3/2}\right).$$

For the i.i.d. case considered by Vervaat [9], i.e. $p_k = p$, all k, this means $nq\to\infty$, $n^2q^3\to 0$, $n\to\infty$, giving

$$(2.31)\qquad d_1(\mu,\nu) = \frac{q}{p}(2\pi e)^{-1/2} + 0\left(\frac{1}{n}\right)$$

$$(2.32)\qquad d_2(\mu,\nu) = \frac{q}{2p}(2\pi e)^{-1/2} + 0\left(\frac{1}{n}\right)$$

and if also $n^3q^5\to 0$, $n\to\infty$,

$$(2.33)\qquad d_3(\mu,\nu) = \frac{1}{2}\sqrt{n}\left(\frac{q}{p}\right)^{3/2}(2\pi)^{-1/2} + 0\left(\sqrt{\frac{q}{n}}\right).$$

Relation (2.31) improves Vervaat's [9] bound

$$(2.34)\qquad d_1(\mu,\nu) \le \frac{q}{p}$$

asymptotically by the factor $(2\pi e)^{-1/2}$. In this low and moderate range for ξ, relations (2.1) to (2.3) will in many cases also improve upon this estimation. For instance, as long as $\xi < 2-\sqrt{2}$, we have

$$(2.35)\qquad \begin{aligned} d_1(\mu,\nu) &\le \min\{\sigma^2(1-\tfrac{1}{2}\xi)e^{-\xi} + 2\sigma^3,\ (2-\sqrt{2})\tfrac{q}{p}\} \\ &\le \min\{.231 + \tfrac{.687}{\sqrt{n}},\ .568\}\,\tfrac{q}{p} < \tfrac{q}{p}\,. \end{aligned}$$

3. Approximation of Mixed Poisson Distributions - the Biased Case

For Poisson approximation with respect to Poisson binomial distributions, it has turned out that within the low range for the mean, unbiased approximation is asymptotically less effective than biased approximation (see Deheuvels and Pfeifer [2], [3]). This suggests that for Poisson approximation in the mixed Poisson case, a similar result should hold true. In fact, the general formulation of Theorem 1 gives an answer to this question.

COROLLARY 3. Let ν be a mixed Poisson distribution with mixing r.v. X such that $E(X) = \zeta > 0$, and μ a Poisson distribution with mean $\xi > 0$. Then

$$(3.1)\qquad \begin{aligned} &|d_i(\mu,\nu) - C_i \,\|(\zeta-\xi)e^{A\xi}Ag_i + \tfrac{1}{2}((\zeta-\xi)^2+\sigma^2)e^{A\xi}A^2g_i\|_{\mathfrak{X}_i}| \\ &\leq \tfrac{1}{3} K_i E(|X-\xi|^3), \text{ if } X \in L^3(P), \end{aligned}$$

where C_i, g_i, $\mathfrak{X}_i$, K_i, $i = 1,2,3$ are as in Corollary 1. If additionally also $X \in L^4(P)$, then the right hand side of (3.1) can also be replaced by (2.3).

A general development for the norm terms in (3.1) can be found in Deheuvels and Pfeifer [2], relation (2.15), and Deheuvels and Pfeifer [3], relations (3.19) and (3.20). We shall not discuss this in detail here but will rather concentrate on the case when $\zeta - \xi = \frac{1}{2}((\zeta-\xi)^2 + \sigma^2)$, or, equivalently,

$$(3.2)\qquad \xi = \zeta - \sigma^2/(1+\sqrt{1-\sigma^2}),$$

provided that $\sigma^2 \leq \min(\zeta,1)$. This choice is asymptotically (when $\sigma^2 \to 0$) optimal in the low range for ζ (provided that $E(|X-\xi|^3) = o(\sigma^2)$), giving

$$(3.3)\qquad d_1(\mu,\nu) = \frac{1}{2}\sigma^2 \frac{\zeta^{[\zeta]}}{[\zeta]!} e^{-\zeta} + o(\sigma^2)$$

$$(3.4)\qquad d_3(\mu,\nu) = \frac{1}{4}\sigma^2 + o(\sigma^2).$$

Surprisingly, for d_2, a biased choice for ξ will in general not result in (essentially) improved estimations; cf. also Theorem 3.1 in Deheuvels and Pfeifer [3]. A biased choice of ξ according to (3.1) will asymptotically (for $\sigma^2 \to 0$) be better than an unbiased choice in (3.3) as long as

$$(3.5)\qquad \zeta < 1 + (\sqrt{2}+1)^{1/3} - (\sqrt{2}-1)^{1/3} \sim 1.596$$

(cf. Deheuvels and Pfeifer [2]), and better than an unbiased choice in

(3.4) as long as

(3.6) $\zeta < \log 2 \sim .693$

(cf. Deheuvels and Pfeifer [3]).

For negative Poisson binomial distributions, we generally have

$$E(X-\xi)^4 = G^{(4)}(0) + 3((\zeta-\xi)^2 + \sigma^2)^2 + 4G'''(0)(\zeta-\xi) + 6((\zeta-\xi)^2+\sigma^2)(\zeta-\xi)^2 + (\zeta-\xi)^4 \tag{3.7}$$

with G as in (2.27), giving

$$E(X-\xi)^4 \leq 14\sigma^4, \text{ if } \sigma^2 \leq 1/4. \tag{3.8}$$

We thus have $E(|X-\xi|^3) = o(\sigma^2)$ for $\sigma^2 \to 0$, hence biased Poisson approximation is asymptotically more effective for negative Poisson distributions in the low range than unbiased approximation. For instance, in the i.i.d. case, under the assumptions that $q \leq .1$, $.04 < \xi < 2-\sqrt{2}$, we have in the biased case

$$d_1(\mu,\nu) \leq \frac{1}{2}\sigma^2(1+\sigma^2)e^{-\xi} + \frac{5}{2}\sigma^3 \leq (.180 + \frac{.859}{\sqrt{n}})\,\frac{q}{p} \tag{3.9}$$

which is asymptotically better than (2.35).

Acknowledgements. We would like to express our deep gratitude to Prof. Stamatis Cambanis for the invitation to the Center for Stochastic Processes, UNC at Chapel Hill, and his warm hospitality which made the stay so pleasant.

References

[1] Barbour, A. D. and Hall, P.: On the rate of Poisson convergence. Math. Proc. Camb. Phil. Soc. 95 (1984), 473 - 480

[2] Deheuvels, P. and Pfeifer, D.: A semigroup approach to Poisson approximation. Ann. Prob. 14 (1986), 663 - 676

[3] Deheuvels, P. and Pfeifer, D.: Operator semigroups and Poisson convergence in selected metrics. Semigroup Forum 34 (1986), 203 - 224

[4] Feller, W.: An Introduction to Probability Theory and Its Applications. II. Wiley, New York, 1971

[5] LeCam, L.: An approximation theorem for the Poisson binomial distribution. Pacific J. Math. 10 (1960), 1181 - 1197

[6] Pfeifer, D.: Probabilistic representations of operator semigroups - a unifying approach. Semigroup Forum 30 (1984), 17 - 34

[7] Pfeifer, D.: Approximation-theoretic aspects of probabilistic representations for operator semigroups. J. Approx. Theory 43 (1985), 271 - 296

[8] Vallender, S. S.: Calculation of the Wasserstein distance between distributions on the line. Theory Prob. Appl. 18 (1973), 784 - 786

[9] Vervaat, W.: Upper bounds for the distance in total variation between the binomial or negative binomial and the Poisson distribution. Statist. Neerlandica 23 (1969), 79 - 86

D. Pfeifer
Institut für Statistik und Wirtschaftsmathematik,
RWTH Aachen, Wüllnerstr. 3, D-5100 Aachen

and

University of North Carolina at Chapel Hill