OPERATOR SEMIGROUPS AND POISSON CONVERGENCE IN SELECTED METRICS

Paul Deheuvels and Dietmar Pfeifer

Communicated by J.A. Goldstein

Introduction

We investigate the approximation of the distribution of the sum of independent (but not necessarily identically distributed) Bernoulli random variables by Poisson distributions with respect to selected probability metrics (total variation, Kolmogorov metric, Fortet-Mourier metric). Special emphasis is given to the problem of the (asymptotically) best choice of the Poisson mean as well as a precise evaluation of the leading term in the corresponding distance. The general approach to these kind of problems chosen here is an appropriate application of the theory of operator semigroups, generalizing and improving earlier results of LeCam (1960).

§ 1. Preliminaries and basic definitions

Let $X_1,..,X_n$ be independent Bernoulli random variables with values in {0,1} and success probabilities

$$p_i(n) = P(X_i = 1), \quad 1 \leq i \leq n$$

(possibly depending on n), and $Y_1,..,Y_n$ be independent Poisson random variables with expectations

$$E(Y_i) = \mu_i(n), \quad 1 \leq i \leq n.$$

We are interested in the approximation of the distribution of $S_n = \sum_{i=1}^{n} X_i$ by the (Poisson) distribution of $T_n = \sum_{i=1}^{n} Y_i$. As a measure of accuracy, we shall consider suitable metrics on the set of all probability measures concentrated on the non-negative integers $\mathbb{Z}^+$, such as

$$(1.1) \quad d(S_n,T_n) = \sup_{M \subseteq \mathbb{Z}^+} \left| P(S_n \in M) - P(T_n \in M) \right| = \frac{1}{2} \sum_{k=0}^{\infty} \left| P(S_n = k) - P(T_n = k) \right|$$

(total variation),

$$(1.2) \quad d_0(S_n,T_n) = \sup_{m \in \mathbb{Z}^+} \left| \sum_{k=0}^{m} \{P(S_n = k) - P(T_n = k)\} \right|$$

(Kolmogorov metric), and

$$(1.3) \qquad d_1(S_n,T_n) = \inf_Q E(|S_n - T_n|) = \sum_{k=0}^{\infty} \left| \sum_{j=0}^{k} \{P(S_n = j) - P(T_n = j)\} \right|$$

where Q ranges through all possible joint distributions $\mathcal{L}(S_n,T_n)$ of (S_n,T_n) with given marginal distributions $\mathcal{L}(S_n)$ and $\mathcal{L}(T_n)$ (Fortet-Mourier metric; a specific Wasserstein metric).

Note that we have used the notation $d(S_n,T_n)$ etc. instead of $d(\mathcal{L}(S_n),\mathcal{L}(T_n))$ for simplicity.
For a good survey on probability metrics the interested reader is referred to Zolotarev (1984) and Vallender (1973). Statistical applications of these metrics are outlined in Serfling (1978) and Deheuvels, Karr, Pfeifer and Serfling (1986).
While the case of Poisson approximation in total variation has received relatively complete treatment by various authors such as LeCam (1960), Kerstan (1964), Chen (1974, 1975), Serfling (1975, 1978), Barbour and Hall (1984) and Deheuvels and Pfeifer (1986 a), relatively little is known with respect to the metrics d_0 and d_1 (see e.g. Franken (1964), Gastwirth (1977), Serfling (1978)). Following an idea developed in Pfeifer (1985 a), we shall show in the sequel that approximation problems of the above kind can very generally be treated in an operator semigroup framework resembling LeCam's (1960)

approach, allowing at the same time also for a discussion on (asymptotically) optimal choice problems for the Poisson mean as in Deheuvels and Pfeifer (1986 a).

For abbreviation, we shall denote by $p(n) = (p_1(n),..,p_n(n))$, $\mu(n) = (\mu_1(n),..,\mu_n(n))$ and $\lambda(n) = (\lambda_1(n),..,\lambda_n(n))$ where

$$(1.4)\qquad \lambda_i(n) = -\log(1-p_i(n)), \qquad 1 \leq i \leq n.$$

The choices $\mu(n) = p(n)$ and $\mu(n) = \lambda(n)$ are of special interest here since they turn out to be asymptotically optimal when

$$(1.5)\qquad \sum_{i=1}^{n} p_i(n) \to \infty\ , \quad \sum_{i=1}^{n} p_i^2(n) = O(1) \qquad (n \to \infty)$$

or

$$(1.6)\qquad \sum_{i=1}^{n} p_i(n) = o(1) \qquad (n \to \infty).$$

More generally, it will be of importance to consider also the choices

$$(1.7)\qquad \mu_i(n) = p_i(n) + \gamma(t)\ p_i^2(n), \qquad 1 \leq i \leq n$$

for some constant $\gamma(t) \in \left[0,\frac{1}{2}\right]$ in case that

$$(1.8)\qquad \sum_{i=1}^{n} p_i(n) \to t,\ \max_{1\leq i\leq n} \{p_i(n)\} = o(1) \qquad (n \to \infty)$$

for some positive finite t. It will be shown in the sequel that under condition (1.8), there will always be such a $\gamma(t)$ which asymptotically minimizes the distance between $\mathcal{L}(S_n)$ and $\mathcal{L}(T_n)$ measured with respect to the above metrics. (The case $t=0$ could have been included here as well since $\lambda_i(n) = p_i(n) + \frac{1}{2} p_i^2(n) + \ldots$ $\ldots + O(p_i^3(n))$, hence $\gamma(0) = \frac{1}{2}$ would be an appropriate choice; correspondingly, under (1.5) which corresponds to the case $t=\infty$, we would have $\gamma(\infty) = 0$.)

§ 2. The semigroup setting of Poisson approximation

Let $\mathfrak{X}$ denote either the Banach space ℓ^1 of all absolutely summable sequences or the Banach space ℓ^∞ of all absolutely bounded sequences $f = (f(0), f(1), \ldots)$, resp. If $\mathcal{M}$ denotes the set of all probability measures concentrated on $\mathbb{Z}^+$, then any measure $m \in \mathcal{M}$ can be identified with the sequence $f_m = (m(\{0\}), m(\{1\}), \ldots)$ $\in \ell^1 \subseteq \mathfrak{X}$. For $f \in \ell^1$, $g \in \mathfrak{X}$ the convolution $f * g$ is defined by

$$(2.1) \qquad f * g(n) = \sum_{k=0}^{n} f(k)\, g(n-k), \qquad n \in \mathbb{Z}^+.$$

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

$$(2.2) \qquad \|f * g\| \leq \|f\|_{\ell^1} \, \|g\|_{\mathfrak{X}} ,$$

where $\|\cdot\|_{\mathfrak{X}}$ denotes the corresponding norm on $\mathfrak{X}$.
Any measure $m \in \mathcal{M}$ thus defines a bounded linear operator on $\mathfrak{X}$ via

$$(2.3) \qquad m\, g = f_m * g = \sum_{k=0}^{\infty} m(\{k\})\, B^k g, \qquad g \in \mathfrak{X}$$

where $Bg = \varepsilon_1 * g$, and ε_k for $k \in \mathbb{Z}^+$ denotes the Dirac measure concentrated at k. If, as usual, I denotes the identity mapping on $\mathfrak{X}$, then $A = B - I$ generates a contraction semigroup given by

$$(2.4) \qquad e^{tA}g = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k g = \sum_{k=0}^{\infty} e^{-t} \frac{t^k}{k!} \varepsilon_k * g = \mathcal{P}(t)\, g, \; g \in \mathfrak{X}$$

for $t \geq 0$, where $\mathcal{P}(t)$ denotes the Poisson distribution with mean t. The semigroup given in (2.4) is also called the Poisson convolution semigroup. It is easy to see that the generator A is a difference operator on $\mathfrak{X}$ which can also be expressed by

$$(2.5) \qquad A\, g(n) = \begin{cases} g(n-1) - g(n), & n \geq 1 \\ -g(0), & n = 0 \end{cases} , \; g \in \mathfrak{X} .$$

On the other hand, if $\mathcal{B}(t_i)$, $1 \leq i \leq n$, denote binomial distributions over $\{0,1\}$ with success probabilities $t_i \in [0,1]$ then also

$$(2.6) \qquad (I + t_i A)g = \mathcal{B}(t_i)g, \quad g \in \mathfrak{X} ,$$

and hence, in the case of independence,

$$(2.7) \qquad \prod_{k=1}^{n} (I + p_i(n)A)g = \mathcal{L}(S_n)\, g, \quad g \in \mathfrak{X} .$$

Further, by (1.1) to (1.3), we see that for measures $m_1, m_2 \in \mathcal{M}$ the metric distances can equivalently be expressed by

(2.8) $$d(m_1,m_2)=\frac{1}{2}\|(m_1-m_2)g\|_{\ell^1} \text{ with } g=(1,0,0,\ldots)$$

(2.9) $$d_0(m_1,m_2)=\|(m_1-m_2)h\|_{\ell^\infty} \text{ with } h=(1,1,1,\ldots)$$

(2.10) $$d_1(m_1,m_2)=\lim_{n\to\infty}\|(m_1-m_2)h_n\|_{\ell^1} \quad \text{with}$$

$h_n = (1,1,..,1,0,0,...)$, containing n ones followed by an infinite string of zeros.

For short, we shall also write

(2.11) $$d_1(m_1,m_2)=\|(m_1-m_2)h\|_{\ell^1} \text{ with h as in (2.9),}$$

although h is not an element of ℓ^1, considering this expression as the limit in (2.10).

With g,h as above, we can thus write

(2.12) $$d(S_n,T_n)=\frac{1}{2}\Big\|(\exp\{\sum_{i=1}^n \mu_i(n)\}A-\prod_{i=1}^n(I+p_i(n))A)g\Big\|_{\ell^1}$$

(2.13) $$d_0(S_n,T_n)=\Big\|(\exp\{\sum_{i=1}^n \mu_i(n)\}A-\prod_{i=1}^n(I+p_i(n))A)h\Big\|_{\ell^\infty}$$

(2.14) $$d_1(S_n,T_n)=\Big\|(\exp\{\sum_{i=1}^n \mu_i(n)\}A-\prod_{i=1}^n(I+p_i(n))A)h\Big\|_{\ell^1}.$$

In order to evaluate these expressions precisely it will be necessary to deal in general with differences

(2.15) $$e^{tA}-\prod_{i=1}^n(I+t_iA) \quad \text{for } t\geq 0,\ t_i\in[0,1]\ .$$

For $t=\sum_{i=1}^n t_i$ we would expect the difference in (2.15) to be close to the operator $\frac{1}{2}\{\sum_{i=1}^n t_i^2\}\, e^{tA}A^2$ in some sense. A more precise statement is as follows.

Theorem 2.1. Let $\mathfrak{X}$ be an arbitrary Banach space and A the bounded generator of a (uniformly continuous) contraction semigroup acting on $\mathfrak{X}$. Let further $t_i \in [0,1]$, $1 \le i \le n$, and $t = \sum_{i=1}^{n} t_i$, $s = \sum_{i=1}^{n} t_i^2$, and $v = \sum_{i=1}^{n} t_i^3$.
Then, for all $f \in \mathfrak{X}$, we have

$$(2.16)\quad \left\| e^{tA}f - \prod_{i=1}^{n}(I + t_iA)\, f - \tfrac{1}{2} s\, e^{tA}A^2 f \right\| \le$$

$$e^{K\|A^2\|s + \|A\|}\left\{ \tfrac{2}{3} v \left\| e^{tA}A^3 f\right\| + \tfrac{K}{2} s^2 \left\| e^{tA}A^4 f\right\| \right\}$$

where $K = 1 + \frac{1}{2}(1 + \|A\|)e^{\|A\|}$.

Proof. We shall make use of a suitable telescoping argument valid for commuting operators $U_1,\ldots,U_n$, $V_1,\ldots,V_n$ on $\mathfrak{X}$ for which we have

$$(2.17)\quad \prod_{i=1}^{n} U_i - \prod_{i=1}^{n} V_i = \sum_{i=1}^{n} U_{i+1}\cdots U_n (U_i - V_i) V_1 \cdots V_{i-1}$$

(cf. also the proof of Theorem 1 in LeCam (1960)).
From this we obtain, letting $w_i = \sum_{j \ne i} t_j$,

$$(2.18)\quad \left\| e^{tA}f - \prod_{i=1}^{n}(I + t_iA)f - \tfrac{1}{2}\sum_{i=1}^{n} t_i^2\, e^{w_iA}A^2 f \right\| \le$$

$$\sum_{i=1}^{n}\Bigg(\prod_{j=1}^{i-1}\left\|e^{-t_jA}(I+t_jA)\right\| \left\|e^{w_iA}\big(e^{t_iA} - (I+t_iA+\tfrac{t_i^2}{2}A^2)\big)f\right\| +$$

$$\ldots + \frac{t_i^2}{2}\left\| e^{w_iA}A^2\left\{\prod_{l=1}^{i-1} e^{-t_lA}(I+t_lA) - I\right\}f\right\| \Bigg).$$

Further, for every $u \geq 0$, we have

$$(2.19) \qquad e^{-uA}(I+uA) = (I+uA)\left\{I - uA + \int_0^u (u-x)e^{-xA}A^2\,dx\right\}$$
$$= I - u^2A^2 + (I+uA)\int_0^u (u-x)e^{-xA}A^2\,dx,$$

hence

$$(2.20) \qquad \left\|e^{-uA}(I+uA)\right\| \leq 1+u^2\|A^2\| + (1+u\|A\|)\frac{u^2}{2}e^{u\|A\|}\|A^2\|$$

which is bounded by $\exp(K\|A^2\|u^2)$ for $u \leq 1$.

Second, we have, by Taylor's formula and the contraction property, for $u,w \geq 0$, $f \in \mathfrak{X}$,

$$(2.21) \qquad \left\|e^{wA}(e^{uA} - (I+uA+\tfrac{u^2}{2}A^2))f\right\| \leq \int_0^u \frac{(u-x)^2}{2}\left\|e^{wA}A^3f\right\|dx$$
$$= \frac{u^3}{6}\left\|e^{wA}A^3f\right\|.$$

Third, by (2.17), we have, for every $w \geq 0$, $f \in \mathfrak{X}$,

$$(2.22) \qquad \left\|e^{wA}A^2\left\{\prod_{l=1}^{i-1} e^{-t_lA}(I+t_lA) - I\right\}f\right\| \leq$$
$$\sum_{l=1}^{i-1}\prod_{j=l+1}^{i-1}\left\|e^{-t_jA}(I+t_jA)\right\|\left\|(e^{-t_lA}(I+t_lA)-I)e^{wA}A^2f\right\|$$
$$= \exp(K\|A^2\|s)\; K\,s\left\|e^{wA}A^4f\right\|$$

where the last estimation follows from (2.19).

Fourth, for $j = 1,2,3$, we have

$$(2.23) \qquad \left\|e^{w_iA}A^jf\right\| \leq \left\|e^{-t_iA}\right\|\left\|e^{tA}A^jf\right\| \leq e^{\|A\|}\left\|e^{tA}A^jf\right\|,$$

hence the right hand side of (2.18) is bounded by

$$(2.24)\qquad e^{K\|A^2\| s + \|A\|}\left\{ \frac{v}{6}\|e^{tA}A^3 f\| + \frac{K}{2} s^2 \| e^{tA}A^4 f\| \right\}.$$

Fifth, we have, again by Taylor's formula,

$$(2.25)\qquad \left\| \sum_{i=1}^{n} t_i^2 \, e^{w_i A} A^2 f - \sum_{i=1}^{n} t_i^2 \, e^{tA} A^2 f \right\| \leq$$

$$\sum_{i=1}^{n} t_i^2 \|(e^{-t_i A} - I) e^{tA} A^2 f\| \leq v\, e^{\|A\|} \| e^{tA} A^3 f \| .$$

The theorem now follows from (2.18), (2.24) and (2.25).∎

It should be pointed out that if $I+uA$, $0 \leq u \leq 1$, also is a contraction (as is the case in our applications, where $I+uA = B(u)$), then K may likewise be replaced by the smaller constant $K^* = 1 + \frac{1}{2} e^{\|A\|}$. Note that in general, the semigroup $\{e^{-uA};\ u \geq 0\}$ used in the proof will in general not be a contraction (as is the case in our applications), but that we only have

$$1 = \|I\| \leq \| e^{-uA} \| \, \| e^{uA} \| \leq \| e^{-uA} \| , \qquad u \geq 0.$$

The following result will be the key for an appropriate treatment of optimal choice problems in Poisson approximation mentioned above.

<u>Theorem 2.2</u>. Under the conditions of Theorem 2.1, we have, for arbitrary $u \geq 0$, $f \in \mathfrak{X}$,

$$(2.26)\qquad \| e^{uA} f - \prod_{i=1}^{n} (I+t_i A) f \| = \frac{s}{2} \| e^{tA}(2\delta A + A^2) f \| + R_n(f)$$

with

$$(2.27)\quad \left|R_n(f)\right| \le e^{K\|A^2\| s+\|A\|}\left\{\frac{2}{3} v\|e^{tA}A^3f\| + \frac{K}{2} s^2\|e^{tA}A^4f\|\right\} + \ldots + 2\delta^2 s^2 \max\left\{\|e^{uA}A^2f\|, \|e^{tA}A^2f\|\right\}$$

where $\delta = \frac{u-t}{s}$.

Proof. We have

$$(2.28)\quad \left\|e^{uA}f - \prod_{i=1}^{n}(I+t_iA)f - \frac{s}{2}e^{tA}(2\delta A+A^2)f\right\| \le$$

$$\left\|e^{tA}f - \prod_{i=1}^{n}(I+t_iA)f - \frac{s}{2}e^{tA}A^2f\right\| + \left\|e^{uA}f - e^{tA}f - (u-t)e^{tA}f\right\|.$$

The assertion now follows by Theorem 2.1 and another application of Taylor's formula as in Deheuvels and Pfeifer (1986 a), Lemma 4.1. ∎

Theorems 2.1 and 2.2 generalize results obtained in Deheuvels and Pfeifer (1986 b) for the special case of Poisson convergence in total variation.

In the sequel we shall apply Theorem 2.2 to the norm representations (2.12) to (2.14) of the various distances under consideration, allowing for precise determination of the leading terms in the asymptotic expansions, and for determination of the asymptotically best choice of the corresponding Poisson parameters. The results obtained here are (at least asymptotically) better than those of LeCam (1960) due to the fact that we consider second order expansions rather than only first order expansions for the semigroup, and that

$\|e^{tA}A\|$ and $\|e^{tA}A^2\|$ tend to zero with t tending to infinity.

§ 3. Consequences for Poisson convergence

Let again denote $g = (1,0,0,\ldots)$, $h = (1,1,1,\ldots)$. Then $Ah = -g$, which is an element of ℓ^1. Since in all investigations, only powers A^k with $k \in \mathbb{N}$ are of interest, no more limit relations have to be considered with respect to asymptotic expansions for the Fortet-Mourier metric (i.e. the right-hand side of (2.26) and (2.27)).

The following expansions and estimations have, in different contexts, partially been obtained earlier (Pfeifer (1985 b), Deheuvels and Pfeifer (1986 a)).

Lemma 3.1. For the Poisson convolution semigroup, we have, for $t \geq 0$,

$$(3.1) \qquad \|e^{tA}Ah\|_{\ell^1} = \sum_{k=0}^{\infty} e^{-t}\frac{t^k}{k!} = 1$$

$$(3.2) \qquad \|e^{tA}A^2h\|_{\ell^1} = \|e^{tA}Ag\|_{\ell^1} = \sum_{k=0}^{\infty} e^{-t}\frac{t^{k-1}}{k!}|t-k| =$$

$$2\,e^{-t}\frac{t^{[\![t]\!]}}{[\![t]\!]!} \sim \frac{2}{\sqrt{2\pi t}} \quad (t\to\infty)$$

where $[\![\,.\,]\!]$ denotes integer part

$$(3.3) \qquad \|e^{tA}A^3h\|_{\ell^1} = \|e^{tA}A^2g\|_{\ell^1} =$$

$$\sum_{k=0}^{\infty} e^{-t}\frac{t^{k-2}}{k!}|t^2-2kt+k(k-1)| =$$

$$2\,e^{-t}\left\{\frac{t^{a-1}(a-t)}{a!}+\frac{t^{b-1}(t-b)}{b!}\right\}\sim\frac{4}{t\sqrt{2\pi e}}\quad(t\to\infty)$$

with $a=\left[\!\left[t+\frac{1}{2}+\sqrt{t+\frac{1}{4}}\right]\!\right]$, $b=\left[\!\left[t+\frac{1}{2}-\sqrt{t+\frac{1}{4}}\right]\!\right]$

(3.4) $$\|e^{tA}A^4h\|_{\ell^1}=\|e^{tA}A^3g\|_{\ell^1}\leq\min\left\{8,\ \frac{6}{t\sqrt{t}}\right\}$$

(3.5) $$\|e^{tA}A^4g\|_{\ell^1}\leq 16\min\left\{1,\ \frac{4}{t^2}\right\}$$

(3.6) $$\|e^{tA}Ah\|_{\ell^\infty}=e^{-t}\sup_{n\geq 0}\frac{t^n}{n!}=e^{-t}\frac{t^{[\![t]\!]}}{[\![t]\!]!}\sim\frac{1}{\sqrt{2\pi t}}\quad(t\to\infty)$$

(3.7) $$\|e^{tA}A^2h\|_{\ell^\infty}=e^{-t}\sup_{n\geq 0}\frac{t^{n-1}}{n!}|t-n|=$$

$$e^{-t}\max\left\{\frac{t^{a-1}(a-t)}{a!},\ \frac{t^{b-1}(t-b)}{b!}\right\}\sim\frac{1}{t\sqrt{2\pi e}}\quad(t\to\infty)$$

(3.8) $$\|e^{tA}A^3h\|_{\ell^\infty}=e^{-t}\sup_{n\geq 0}\frac{t^{n-2}}{n!}|t^2-2nt+n(n-1)|\leq$$

$$\min\left\{2,\ \frac{5}{t\sqrt{t}}\right\}$$

(3.9) $$\|e^{tA}A^4h\|_{\ell^\infty}\leq 16\min\left\{1,\ \frac{4}{t^2}\right\}.$$

Further, by (2.5), we see that in any case, we have $\|A\|\leq 2$, hence $K=13$ ($K^*=5$, resp.) would also be an admissible value in (2.16) and (2.27).

A simple comparison of Lemma 3.1 and Theorem 2.2 with relations (2.12) to (2.14) shows that under condition (1.8) the asymptotically optimal choice for the Poisson mean $\mu=\sum_{i=1}^{n}\mu_i(n)$ is actually given by (1.7), at least for some real constant $\gamma(t)$. The following expansions show that we may indeed restrict the range for $\gamma(t)$

to the interval $\left[0,\frac{1}{2}\right]$.

Lemma 3.2. For the Poisson convolution semigroup, we have, for $t \geq 0$ and every real δ,

$$(3.10)\quad \|e^{tA}(2\delta A + A^2)g\|_{\ell^1} =$$

$$e^{-t}\sum_{k=0}^{\infty}\frac{t^{k-2}}{k!}\left|k^2 - 2k(t+\tfrac{1}{2}-\delta t) + t(t-2\delta t)\right| =$$

$$2\,e^{-t}\left\{\frac{t^{c-1}(c-(1-2\delta)t)}{c!} + \frac{t^{d-1}((1-2\delta)t-d)}{d!}\right\}$$

$$\sim \frac{2}{t\sqrt{2\pi}}\left\{\zeta\exp\left(-\frac{1}{2\zeta^2}\right) + \frac{1}{\zeta}\exp\left(-\frac{1}{2}\zeta^2\right)\right\}$$

$$\geq \frac{4}{t\sqrt{2\pi e}} \qquad (t \to \infty)$$

where $c = \left[\!\left[t - r + \sqrt{t+r^2}\right]\!\right]$, $d = \left[\!\left[t - r - \sqrt{t+r^2}\right]\!\right]$ and $r = \delta t - \frac{1}{2}$, $\zeta = \delta\sqrt{t} + \sqrt{1+\delta^2 t}$

$$(3.11)\quad \|e^{tA}(2\delta A+A^2)h\|_{\ell^\infty} = e^{-t}\sup_{n\geq 0}\frac{t^{n-1}}{n!}\left|(1-2\delta)t - n\right| =$$

$$e^{-t}\max\left\{\frac{t^{c-1}(c-(1-2\delta)t)}{c!},\ \frac{t^{d-1}((1-2\delta)t-d)}{d!}\right\}$$

$$\sim \frac{1}{t\sqrt{2\pi}}\max\left\{\zeta\exp\left(-\frac{1}{2\zeta^2}\right),\ \frac{1}{\zeta}\exp\left(-\frac{1}{2}\zeta^2\right)\right\}$$

$$\sim \frac{\zeta}{t\sqrt{2\pi}}\exp\left(-\frac{1}{2\zeta^2}\right) \qquad (t \to \infty)$$

$$(3.12)\quad \|e^{tA}(2\delta A+A^2)h\|_{\ell^1} = e^{-t}\sum_{k=0}^{\infty}\frac{t^{k-1}}{k!}\left|k - (1-2\delta)t\right| =$$

$$2\delta - 4\delta\sum_{k=0}^{N} e^{-t}\frac{t^k}{k!} + 2\,e^{-t}\frac{t^N}{N!}$$

$$\sim \frac{2}{\sqrt{2\pi(1-2\delta)t}} \exp\left(-\frac{2\delta^2 t}{1-2\delta}\right) + 2\delta - 4\delta\,\phi(-2\delta\sqrt{t}) \quad (t\to\infty)$$

where $N = [\![(1-2\delta)t]\!]$ and $\phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du$,

and for $\delta > 0$,

$$4\delta\phi(-2\delta\sqrt{t}) \sim \frac{2}{\sqrt{2\pi t}} \exp(-2\delta^2 t) \qquad (t\to\infty).$$

Note that some of the above asymptotic relations are valid only for $\delta \geq 0$, but from the exact representation of the norm terms it is easy to see that only this range for δ is the interesting one.

Correspondingly, we see that under condition (1.6), in all three cases the choice $\delta = \frac{1}{2}$ gives the asymptotically minimal value for the norm terms above. The fact that instead of taking $\mu_i(n) = p_i(n) + \frac{1}{2}\, p_i^2(n)$, we prefer the choice $\mu(n) = \lambda(n)$ (which is asymptotically equivalent) is due to the circumstance that it can be shown by different methods (see Serfling (1975), (1978) and Deheuvels and Pfeifer (1986 a)) that the latter is precisely optimal in some cases. Intuitively, we should expect that under (1.6), the difference between the probabilities of observing zero will be the major contribution to the metric distances above, which is

$$(3.13) \qquad \left| \prod_{k=1}^{n} (1 - p_k(n)) - \exp\left\{\sum_{k=1}^{n} \mu_k(n)\right\} \right|$$

and which is zero for the choice $\mu(n) = \lambda(n)$.

In what follows we shall for short omit the variable n in $p(n)$, $\mu(n)$ etc., writing p, μ etc., and we shall write $T_n(\mu)$ instead of T_n, expressing that T_n follows a Poisson distribution with mean $\sum_{i=1}^{n}\mu_i$.

The following statement gives a summary of the results, without explicitly appealing to the operator-theoretic background.

Theorem 3.1.

A) Under the condition $\sum_{i=1}^{n} p_i \to 0$ $(n \to \infty)$ we have

(3.14) $\inf_{\mu} d(S_n, T_n(\mu)) \sim d(S_n, T_n(\lambda)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2$

(3.15) $\inf_{\mu} d_0(S_n, T_n(\mu)) \sim d_0(S_n, T_n(\lambda)) \sim d_0(S_n, T_n(p))$

$\sim \frac{1}{2} \sum_{i=1}^{n} p_i^2$

(3.16) $\inf_{\mu} d_1(S_n, T_n(\mu)) \sim d_1(S_n, T_n(\lambda)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2$

(3.17) $d(S_n, T_n(p)) \sim d_1(S_n, T_n(p)) \sim \sum_{i=1}^{n} p_i^2$,

all relations valid for $n \to \infty$.

B) Under the condition $\sum_{i=1}^{n} p_i \to t \in (0, \infty)$, $\max_{1 \le i \le n}\{p_i\} \to 0$ $(n \to \infty)$, we have

(3.18) $d(S_n, T_n(p)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2 \; e^{-t}\left\{\frac{t^{a-1}(a-t)}{a!} + \frac{t^{b-1}(t-b)}{b!}\right\}$

with a,b as in (3.3)

(3.19) $d(S_n, T_n(\lambda)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2 \; e^{-t} \frac{t^{[\![t]\!]}}{[\![t]\!]!}$

which is asymptotically larger than $d(S_n, T_n(p))$ iff $t > 1 + (\sqrt{2}+1)^{1/3} - (\sqrt{2}-1)^{1/3} = 1.5960\ldots$

(3.20) $$d_0(S_n, T_n(p)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2 \; e^{-t} \max\left\{ \frac{t^{a-1}(a-t)}{a!}, \frac{t^{b-1}(t-b)}{b!} \right\}$$

with a,b as in (3.3)

(3.21) $$d_0(S_n, T_n(\lambda)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2 \; e^{-t} \frac{t^{[\![t]\!]}}{[\![t]\!]!}$$

which is asymptotically always larger than $d_0(S_n, T_n(p))$

(3.22) $$d_1(S_n, T_n(p)) \sim \sum_{i=1}^{n} p_i^2 \; e^{-t} \frac{t^{[\![t]\!]}}{[\![t]\!]!}$$

(3.23) $$d_1(S_n, T_n(\lambda)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2$$

which is asymptotically larger than $d_1(S_n, T_n(p))$ iff $t > \log 2 = .6931\ldots$,

all relations valid for $n \to \infty$.

C) Under the condition $\sum_{i=1}^{n} p_i \to \infty$, $\sum_{i=1}^{n} p_i^2 = O(1)$ $(n \to \infty)$

we have

(3.24) $$\inf_{\mu} d(S_n, T_n(\mu)) \sim d(S_n, T_n(p)) \sim \frac{1}{\sqrt{2\pi e}} \frac{\sum_{i=1}^{n} p_i^2}{\sum_{i=1}^{n} p_i}$$

(3.25) $$d(S_n, T_n(\lambda)) \sim \frac{1}{2\sqrt{2\pi}} \frac{\sum_{i=1}^{n} p_i^2}{\sqrt{\sum_{i=1}^{n} p_i}}$$

$$(3.26)\quad \inf_{\mu} d_0(S_n, T_n(\mu)) \sim d_0(S_n, T_n(p))$$

$$\sim \frac{1}{2\sqrt{2\pi e}} \frac{\sum_{i=1}^{n} p_i^2}{\sum_{i=1}^{n} p_i}$$

$$(3.27)\quad d_0(S_n, T_n(\lambda)) \sim \frac{1}{2\sqrt{2\pi}} \frac{\sum_{i=1}^{n} p_i^2}{\sqrt{\sum_{i=1}^{n} p_i}}$$

$$(3.28)\quad \inf_{\mu} d_1(S_n, T_n(\mu)) \sim d_1(S_n, T_n(p)) \sim \frac{1}{\sqrt{2\pi}} \frac{\sum_{i=1}^{n} p_i^2}{\sqrt{\sum_{i=1}^{n} p_i}}$$

$$(3.29)\quad d_1(S_n, T_n(\lambda)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2 ,$$

all relations valid for $n \to \infty$.

Furthermore, under the conditions specified in B), in all three cases, there will always be an asymptotically optimal choice for $\mu = \mu(n)$ of the form

$$(3.30)\quad \mu_i(n) = p_i(n) + \gamma(t)\, p_i^2(n), \quad 1 \leq i \leq n$$

for some constant $\gamma(t) \in [0, \frac{1}{2}]$ which can be determined by the corresponding minimal value of δ in Lemma 3.2.

It should be pointed out that there is in general not a nice closed expression for $\gamma(t)$. Some examples are given below.

<u>Example 3.1</u>.

I) For the total variation distance, we have

(3.31) If $0 < t \leq 1$: $\gamma(t) = \frac{1}{2}$,

$$d(S_n, T_n(\mu)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2 \ e^{-t}$$

(3.32) If $1 < t \leq \sqrt{2}$: $\gamma(t) = \frac{1}{2}$,

$$d(S_n, T_n(\mu)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2 \ t \ e^{-t}$$

(3.33) If $\sqrt{2} < t \leq \sqrt[3]{6}$: $\gamma(t) = \frac{1}{2} - \frac{3}{2t} \frac{2-t}{3-t}$,

$$d(S_n, T_n(\mu)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2 \ e^{-t} \left\{ t + (1 - \frac{t^2}{2}) \frac{3}{t} \frac{2-t}{3-t} \right\}$$

(3.34) If $\sqrt[3]{6} < t \leq 2$: $\gamma(t) = 0$,

$$d(S_n, T_n(\mu)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2 \ e^{-t} \left\{ \frac{t^2}{2} + (1 - \frac{t^3}{6}) \right\}.$$

II) For the Kolmogorov distance, we have

(3.35) If $0 < t \leq \sqrt{3} - 1$: $\gamma(t) = \frac{t}{2(1+t)} = \frac{1}{2} - \frac{1}{2(1+t)}$,

$$d_0(S_n, T_n(\mu)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2 \ \frac{1}{1+t} \ e^{-t}$$

(3.36) If $\sqrt{3} - 1 < t \leq 1$: $\gamma(t) = \frac{1}{2} - \frac{t}{2+t^2}$,

$$d_0(S_n, T_n(\mu)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2 \ \frac{2t}{2+t^2} \ e^{-t}.$$

III) For the Fortet-Mourier distance, we have

(3.37) If $0 < t \leq \log 2$: $\gamma(t) = \frac{1}{2}$,

$$d_1(S_n, T_n(\mu)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2$$

(3.38) If $\log 2 < t \leq 1$: $\gamma(t) = 0$,

$$d_1(S_n, T_n(\mu)) \sim \sum_{i=1}^{n} p_i^2 \; e^{-t}$$

(3.39) If $1 < t \leq \alpha = 1.6784\ldots$: $\gamma(t) = \frac{1}{2} - \frac{1}{2t}$,

$$d_1(S_n, T_n(\mu)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2 \left\{1 - \frac{1}{t}(1-2e^{-t})\right\}$$

(3.40) If $\alpha < t \leq 2$: $\gamma(t) = 0$,

$$d_1(S_n, T_n(\mu)) \sim \sum_{i=1}^{n} p_i^2 \; t \; e^{-t}.$$

Here α denotes the positive root of the equation $2\,e^{-\alpha}(1+\alpha) = 1$.

Finally, we should like to mention that besides asymptotic expansions as given in Theorem 3.1, Theorem 2.2 also allows for upper and lower bounds for the metric distances under consideration, in connection with Lemma 3.1.

Acknowledgements. The second author is indebted to Deutsche Forschungsgemeinschaft for partial financial support by a Heisenberg research grant. Also, we would like to thank J.A. Goldstein for his careful reading of the manuscript.

References

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

Chen, L.H.Y. (1974): On the convergence of Poisson binomial to Poisson distributions. Ann.Prob. 2, 178 - 180.

Chen, L.H.Y. (1975): Poisson approximation for dependent trials. Ann.Prob. 3, 534 - 545.

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

Deheuvels, P. and Pfeifer, D. (1986 b): Semigroups and Poisson approximation. In: Perspectives and New Directions In Theoretical And Applied Statistics, M. Puri, J.P. Villaplana, W. Wertz, eds., Wiley, N.Y. (to appear).

Deheuvels, P., Karr, A.F., Pfeifer, D. and Serfling, R.J. (1986): Poisson approximation in selected metrics by coupling and semigroup methods, with statistical applications. Technical Report, RWTH Aachen (unpublished).

Franken, P. (1964): Approximation der Verteilungen von Summen unabhängiger nichtnegativer ganzzahliger Zufallsgrößen durch Poissonsche Verteilungen. Math.Nachr. 27, 303 - 340.

Gastwirth, J.L. (1977): A probability model of a pyramid scheme. Amer.Statist. 31, 79 - 82.

Kerstan, J. (1964): Verallgemeinerung eines Satzes von Prochrow und LeCam. Z.Wahrscheinlichkeitsth. 2, 173 - 179.

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

Pfeifer, D. (1985 a): A semigroup setting for distance measures in connexion with Poisson approximation. Semigroup Forum 31, 201 - 205.

Pfeifer, D. (1985b): On the distance between mixed Poisson and Poisson distributions. Technical Report No. 115, Center for Stochastic Processes, UNC at Chapel Hill.

Serfling, R.J. (1975): A general Poisson approximation theorem. Ann.Prob. 3, 726 - 731.

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

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

Zolotarev, V.M. (1984): Probability metrics. Th.Prob. Appl. 28, 278 - 302.

Paul Deheuvels
Laboratoire de Statistique
Théorique et Appliquée
Université Pierre et Marie Curie
4 Place Jussieu
75252 Paris Cedex 05
FRANCE

Dietmar Pfeifer
Institut für Statistik
RWTH Aachen
Wüllnerstr. 3
5100 Aachen
GERMANY