Ann. Inst. Statist. Math.
Vol. 40, No. 4, 671–681 (1988)

ON A RELATIONSHIP BETWEEN USPENSKY'S THEOREM AND POISSON APPROXIMATIONS

P. DEHEUVELS[1] AND D. PFEIFER[2]

[1] *Université Paris VI, 7 Avenue du Château, F-92340 Bourg-La-Reine, France*
[2] *Mathematics, University of Oldenburg, Postfach 25 03, D-2900 Oldenburg, West Germany*

(Received January 6, 1987; revised November 18, 1987)

Abstract. In this paper we show that Uspensky's expansion theorem for the Poisson approximation of the distribution of sums of independent Bernoulli random variables can be rewritten in terms of the Poisson convolution semigroup. This gives rise to exact evaluations and simple remainder term estimations for the deviations of the distributions in study with respect to various probability metrics, generalizing results of Shorgin (1977, *Theory Probab. Appl.*, **22**, 846–850). Finally, we compare the sharpness of Poisson versus normal approximations.

Key words and phrases: Poisson approximation, Uspensky's theorem, operator semigroups, probability metrics.

1. Introduction

There has been a considerable recent interest in operator methods in connection with Poisson approximation problems (cf. Pfeifer (1983, 1985), Barbour and Hall (1984), Deheuvels and Pfeifer (1986*a*, 1986*b*, 1987), Barbour (1987)), using e.g. Stein's method (Barbour and Hall (1984), Barbour (1987)) or a Poisson convolution semigroup approach on a suitable Banach space, generalizing an operator-theoretic approach originally due to LeCam (1960). The latter approach has the advantage of being applicable simultaneously to Poisson approximation problems in different probability metrics (cf. Zolotarev (1984)), such as total variation, Kolmogorov distance or Fortet-Mourier distance, and others (see Deheuvels and Pfeifer (1987) for details). Also, it allows for the treatment of (asymptotically) optimal choice problems for the Poisson mean (Deheuvels and Pfeifer (1986*a*, 1986*b*, 1987)). On the contrary, Stein's method has the advantage to be applicable also to other than 0-1-valued random variables (see Barbour (1987)). However, it seems that a very early paper of Uspensky (1931) has been overlooked by most authors dealing with such approximation problems (see, e.g., Serfling (1978) for a survey and references).

Although Uspensky's work is based on complex analysis, it is nevertheless possible to establish a simple relationship between his expansion theorem and the forementioned Poisson convolution semigroup, which simplifies considerably remainder term estimations as in Shorgin (1977) (who considers Poisson approximation in Kolmogorov distance only), and allows at the same time for an extension of the results also to other probability metrics, as such above. It is thus possible to give a complete treatment of Poisson approximation problems of this kind by combination, covering also the cases which remained open by the former approach (see Deheuvels and Pfeifer (1986*a*, 1986*b*, 1987)).

2. Poisson approximation in selected metrics

Let $X_1,\dots,X_n$ be independent Bernoulli random variables with values $\{0,1\}$ with success probabilities

$$p_i(n)=P(X_i=1),\qquad 1\le i\le n, \tag{2.1}$$

(possibly depending on n), and $Y_1,\dots,Y_n$ be independent Poisson random variables with means

$$E(Y_i)=\mu_i(n),\qquad 1\le i\le n. \tag{2.2}$$

Let $S_n=\sum_{i=1}^n X_i$, $T_n=\sum_{i=1}^n Y_i$. We are interested in the approximation of the distribution $\mathscr{L}(S_n)$ of S_n by the Poisson law $\mathscr{L}(T_n)$, with respect to different probability metrics such as

$$\begin{aligned} d(\mathscr{L}(S_n),\mathscr{L}(T_n)) &= \sup_{M\subseteq Z^+}|P(S_n\in M)-P(T_n\in M)| \\ &= \frac{1}{2}\sum_{k=0}^{\infty}|P(S_n=k)-P(T_n=k)|, \end{aligned} \tag{2.3}$$

(total variation),

$$d_0(\mathscr{L}(S_n),\mathscr{L}(T_n))=\sup_x|F_{S_n}(x)-F_{T_n}(x)|, \tag{2.4}$$

(Kolmogorov distance) where F_X denotes the cumulative distribution function of X,

$$d_1(\mathscr{L}(S_n),\mathscr{L}(T_n))=\inf_Q E|S_n-T_n|=\sum_{k=0}^{\infty}|F_{S_n}(k)-F_{T_n}(k)|, \tag{2.5}$$

(Fortet-Mourier distance) where Q ranges through all joint distributions $\mathscr{L}(S_n, T_n)$ with given marginals $\mathscr{L}(S_n)$ and $\mathscr{L}(T_n)$.

Asymptotic evaluations of these distances, combined with suitable remainder term estimations, as well as (asymptotically) optimal choice problems for $\mu_i(n)$ have been treated completely in Deheuvels and Pfeifer (1986*a*, 1986*b*, 1987), for the case that

$$\sum_{i=1}^{n} p_i^2(n) = O(1) \quad (n \to \infty) . \tag{2.6}$$

Especially, it has been shown that if

$$\sum_{i=1}^{n} p_i(n) \to t \in [0, \infty] \quad \text{and} \quad \max_i \{p_i(n)\} \to 0, \quad (n \to \infty) , \tag{2.7}$$

then there exists a real number $\gamma(t) \in [0, 1/2]$ (depending upon the metrics under consideration only) such that asymptotically,

$$\mu_i(n) = p_i(n) + \gamma(t) p_i^2(n), \quad 1 \le i \le n , \tag{2.8}$$

minimizes the distance between $\mathscr{L}(S_n)$ and $\mathscr{L}(T_n)$ for $n \to \infty$. It should be noted that the second condition of (2.7) is not necessary for the case $t = \infty$.

In particular, in all three choices considered above, we may take

$$\gamma(0) = \frac{1}{2}, \quad \gamma(\infty) = 0 , \tag{2.9}$$

whereas in general, $\gamma(t)$ are more complicated functions of t in the range $0 < t < \infty$, differing also for different metrics (see Deheuvels and Pfeifer (1987) for detailed evaluations). Especially, for $t = \infty$, it was shown that with the (asymptotically) optimal choice $\mu_i(n) = p_i(n)$, we have

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

$$d_0(\mathscr{L}(S_n), \mathscr{L}(T_n)) \sim \frac{1}{2} d(S_n, T_n) \sim \frac{1}{2\sqrt{2\pi e}} \frac{\sum_{n=1}^{n} p_i^2(n)}{\sum_{n=1}^{n} p_i(n)} , \tag{2.10}$$

$$d_1(\mathscr{L}(S_n), \mathscr{L}(T_n)) \sim \frac{1}{\sqrt{2\pi}} \frac{\sum_{n=1}^{n} p_i^2(n)}{\sqrt{\sum_{n=1}^{n} p_i(n)}} \quad (n \to \infty) .$$

In this paper, we want to show that these relations are generally valid under the sole conditions

$$\sum_{i=1}^{n} p_i^2(n) = o\left(\sum_{i=1}^{n} p_i(n)\right) \quad (n \to \infty), \tag{2.11}$$

for d and d_0, and

$$\sum_{i=1}^{n} p_i^2(n) = o\left(\sqrt{\sum_{i=1}^{n} p_i(n)}\right) \quad (n \to \infty),$$

for d_1. This improvement is mainly due to a suitable adoption of Uspensky's (1931) and Shorgin's ideas (1977) stemming from complex analysis, in combination with the former semigroup technique. Conditions (2.11) turn out to be necessary for Poisson convergence in the metrics under consideration (see Section 4 for further details concerning d_0), and thus complete the investigations of Poisson approximation problems for Poisson-Bernoulli distributions $\mathscr{L}(S_n)$. Besides evaluations of the leading terms in (2.10) we also obtain simple estimations of the remainder terms, even up to an arbitrary order. These are compared with Barbour's recent results (1987).

3. The Poisson convolution semigroup

Let as usual denote l^1 the Banach space of all absolutely summable sequences $f = (f(0), f(1), \ldots)$. For $f, g \in l^1$, the convolution $f * g$ is defined by

$$f * g(n) = \sum_{k=0}^{n} f(k) g(n-k), \qquad n \geq 0. \tag{3.1}$$

Then $f * g \in l^1$ again, with

$$\|f * g\| \leq \|f\| \, \|g\|. \tag{3.2}$$

Any such element f hence defines a bounded linear operator on l^1 with convolution as operation, with corresponding operator norm identical with $\|f\|$. For $f = (0, 1, 0, 0, \ldots)$ let B denote the corresponding (shift) operator, and I denote the identity (the measures concentrated in $\boldsymbol{Z}^+$ will be identified by the corresponding (summable) sequence of individual probabilities). The bounded operator $A = B - I$ generates a contraction semigroup on l^1 given by

$$e^{tA} g = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k g = \sum_{k=0}^{\infty} e^{-t} \frac{t^k}{k!} B^k g = \mathscr{P}(t) g, \qquad g \in l^1, \tag{3.3}$$

where $\mathscr{P}(t)$ denotes the Poisson distribution with mean t ($t \geq 0$). This semigroup is also called the Poisson convolution semigroup. Its importance is given by the fact that by our assumptions, $\mathscr{L}(S_n)$ corresponds to the operator $\prod_{i=1}^{n} (I + p_i(n)A)$ which is close to $\exp\left(\sum_{i=1}^{n} p_i(n)A\right)$ up to a term roughly like $(1/2)\left(\sum_{i=1}^{n} p_i^2(n)\right) A^2 \exp\left(\sum_{i=1}^{n} p_i(n)A\right)$ such that e.g., under (2.6), when $\sum_{i=1}^{n} p_i(n) \to \infty$, $(n \to \infty)$,

$$d(\mathscr{L}(S_n), \mathscr{L}(T_n)) \sim \frac{1}{4} \sum_{i=1}^{n} p_i^2(n) \left\| A^2 \exp\left(\sum_{i=1}^{n} p_i(n)\right) A \right\| , \tag{3.4}$$

$$d_1(\mathscr{L}(S_n), \mathscr{L}(T_n)) \sim \frac{1}{2} \sum_{i=1}^{n} p_i^2(n) \left\| A \exp\left(\sum_{i=1}^{n} p_i(n)\right) A \right\| . \tag{3.5}$$

By changing the underlying Banach space to the space l^∞ of absolutely bounded sequences, a similar evaluation is possible also for $d_0(\mathscr{L}(S_n), \mathscr{L}(T_n))$ (see Deheuvels and Pfeifer (1986*b*)). It was also proved in Deheuvels and Pfeifer (1987) (cf. also Deheuvels and Pfeifer (1986*a*)) that for all $t > 0$,

$$\|Ae^{tA}\| = \sum_{k=0}^{\infty} e^{-t} \frac{t^{k-1}}{k!} |t-k| = 2e^{-t} \frac{t^{[t]}}{[t]!} \sim \frac{2}{\sqrt{2\pi e}} \quad (t \to \infty) , \tag{3.6}$$

$$\begin{aligned} \|A^2 e^{tA}\| &= \sum_{k=0}^{\infty} e^{-t} \frac{t^{k-2}}{k!} |t^2 - 2kt + k(k-1)| \\ &= 2e^{-t} \left\{ \frac{t^{a-1}(a-t)}{a!} + \frac{t^{b-1}(b-t)}{b!} \right\} \sim \frac{4}{t\sqrt{2\pi e}} \quad (t \to \infty) , \end{aligned} \tag{3.7}$$

where $a = [t + 1/2 + \sqrt{t + 1/4}]$, $b = [t + 1/2 - \sqrt{t + 1/4}]$, and $[\cdot]$ denotes integer part, which gives rise to (2.10) (similar for d_0). Further, it can be seen from (3.6) and (3.7) that we have the bound

$$\|Ae^{tA}\| \leq \sqrt{\frac{2}{et}} \quad (t > 0) , \tag{3.8}$$

which is reached for $t = 1/2$.

The following is a reformulation of Uspensky's theorem (1931), in terms of the Poisson convolution semigroup, for the case of Poisson-Bernoulli distributions.

THEOREM 3.1. *For all $n \in \boldsymbol{N}$, $k \in \boldsymbol{Z}^+$, we have*

(3.9) $$F_{S_n}(k) - F_{T_n}(k) = \frac{\lambda_2}{2} Ae^{\lambda A}g(k) + \left\{ \sum_{j=3}^{\infty} a_j(-A)^{j-1}e^{\lambda A} \right\} g(k) ,$$

(3.10) $$P(S_n = k) - P(T_n = k) = -\frac{\lambda_2}{2} A^2e^{\lambda A}g(k) + \left\{ \sum_{j=3}^{\infty} a_j(-A)^{j}e^{\lambda A} \right\} g(k) ,$$

where $g = (1, 0, 0, \ldots) \in l^1$, $\lambda = \sum_{i=1}^{n} p_i(n)$, $\lambda_m = \sum_{i=1}^{n} p_i^m(n)$ $(m \geq 2)$, *and*

(3.11) $$a_3 = -\frac{\lambda_3}{3}, \quad a_j = -\frac{1}{j}\left(\lambda_j + \sum_{i=2}^{j-2} a_i\lambda_{j-i} \right), \quad j \geq 4 .$$

PROOF. (3.9) is a direct consequence of Uspensky's theorem (cf. Shorgin (1977), Lemma 1). To derive (3.10) from (3.9) observe that for all $k \geq 1$,

$$P(S_n = k) - P(T_n = k) = -A[F_{S_n}(\cdot) - F_{T_n}(\cdot)](k) ;$$

likewise for $k = 0$.

COROLLARY 3.1. *Let* $\theta = \lambda_2/\lambda$. *Then for all n,*

(3.12) $$d(\mathscr{L}(S_n), \mathscr{L}(T_n)) = \frac{\lambda_2}{4} \|A^2e^{\lambda A}\| + r ,$$

with

$$|r| \leq \frac{1}{2} \sum_{j=3}^{\infty} \sqrt{2\theta}^{\,j} = \frac{1}{2} \frac{(2\theta)^{3/2}}{1 - \sqrt{2\theta}} \quad \left(\theta < \frac{1}{2} \right),$$

(3.13) $$d_1(\mathscr{L}(S_n), \mathscr{L}(T_n)) = \frac{\lambda_2}{2} \|Ae^{\lambda A}\| + r_1 ,$$

with

$$|r_1| \leq \sqrt{\lambda_2} \sum_{j=2}^{\infty} \sqrt{2\theta}^{\,j} = 2 \frac{\lambda_2}{\sqrt{\lambda}} \frac{\lambda^{-5/8}}{1 - \sqrt{2\theta}} \quad \left(\theta < \frac{1}{2} \right).$$

PROOF. By Theorem 3.1 and (2.3),

$$|r| \leq \frac{1}{2} \sum_{j=3}^{\infty} |a_j| \, \|A^je^{\lambda A}\| ;$$

but

$$|a_j| \le \left(\frac{e}{j} \lambda_2 \right)^{j/2} = b_j , \tag{3.14}$$

say (Shorgin (1977), Lemma 5), and by (3.8),

$$\|A^j e^{\lambda A}\| \le \|A e^{\lambda/jA}\|^j \le \left(\frac{2j}{e\lambda} \right)^{j/2} , \tag{3.15}$$

which gives (3.12). By (2.5) and (3.9),

$$|r_1| \le \sum_{j=2}^{\infty} |a_{j+1}| \, \|A^j e^{\lambda A}\| \le \sqrt{\lambda_2} \sum_{j=2}^{\infty} b_j \|A^j e^{\lambda A}\| ,$$

which gives (3.13).

Although it is possible to derive similar estimations for $d_0(\mathscr{L}(S_n), \mathscr{L}(T_n))$ by means of the Poisson convolution semigroup on l^∞, we shall omit this here since Shorgin's results (1977) are equally precise for this case. From his Theorem 1, we obtain

$$d_0(\mathscr{L}(S_n), \mathscr{L}(T_n)) = \frac{\lambda_2}{2} e^{-\lambda} \max\left\{ \frac{\lambda^{a-1}(a-\lambda)}{a!} , \frac{\lambda^{b-1}(\lambda - b)}{b!} \right\} + r_0 , \tag{3.16}$$

with

$$|r_0| \le \frac{5}{3} \left\{ \frac{\lambda_3}{\lambda\sqrt{\lambda}} + \frac{\theta^2}{1 - \sqrt{\Theta}} \right\} \quad (\theta < 1) ,$$

where the leading term is asymptotically $\lambda_2 / 2\lambda\sqrt{2\pi e}$ for $n \to \infty$, under (2.11).

Here, a and b are as in (3.7) for $t = \lambda$, which shows why asymptotically $d_0(\mathscr{L}(S_n), \mathscr{L}(T_n)) \sim (1/2) d(\mathscr{L}(S_n), \mathscr{L}(T_n))$ for $n \to \infty$ in this case.

From Corollary 3.1 above it is immediately clear that the conditions specified in (2.11) are necessary and sufficient for Poisson convergence in the metrics considered, and yield the asymptotic expansions given in (2.10).

Of course, from what has been said above it is obvious that expansions of the form (2.10) can be given up to any desired accuracy. For instance, in the case of total variation, we have likewise, for any $m \ge 2$,

$$d(\mathscr{L}(S_n), \mathscr{L}(T_n)) = \frac{1}{2} \left\| \sum_{j=2}^{m} a_j(-A)^j e^{\lambda A} \right\| + r(m) , \tag{3.17}$$

with $a_2 = -\lambda_2/2$, and

$$|r(m)| \le \frac{1}{2}\frac{\sqrt{2\theta}^{m+1}}{1-\sqrt{2\theta}} \qquad \left(\theta < \frac{1}{2}\right). \tag{3.18}$$

Here,

$$\left\| \sum_{j=2}^{m} a_j(-A)^j e^{\lambda A} \right\| = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}\left| \sum_{j=2}^{m}(-1)^j a_j \sum_{i=0}^{j}(-1)^i \binom{j}{i}\frac{k_{(i)}}{\lambda^i} \right| \tag{3.19}$$
$$= \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}\left| \sum_{j=2}^{m} a_j C_j(\lambda, k) \right|,$$

where $C_j(\lambda, k)$ is a Charlier polynomial as in Barbour ((1987), (2.7)). This expression is in general not easily evaluable in closed form, except for $m = 2$ as has been done above.

Remark. A more elaborate evaluation of the total variation distance, considering terms up to a_4, gives

$$d(\mathscr{L}(S_n), \mathscr{L}(T_n)) = \frac{\lambda_2}{4}\,\|A^2 e^{\lambda A}\| + r^*, \tag{3.20}$$

with

$$|r^*| \le .55\,\frac{\lambda_3}{\lambda\sqrt{\lambda}} + .55\,\theta^2 + 1.09\,\frac{\lambda_4}{\lambda^2} + 2.83\,\frac{\theta^{5/2}}{1-\sqrt{2\theta}} \qquad \left(\theta < \frac{1}{2}\right), \tag{3.21}$$

which is asymptotically sharper than the bound in Barbour and Hall (1984). Comparison with Barbour's results (1987) shows that in the range where λ is large the above formulae are more useful, while when λ remains bounded we need higher order terms in order to get the same order of accuracy (e.g., if $p_i(n) \sim \lambda/n$, then (3.19) requires terms of order n^{-m+1} to achieve an error of order $n^{-(m+1)/2}$).

4. Normal approximations

Let $X_1, \dots, X_n$ again be independent Bernoulli random variables such that (2.1) holds and set $S_n = \sum_{i=1}^{n} X_i$, $\lambda = E(S_n) = \sum_{i=1}^{n} p_i(n)$, $\lambda_m = \sum_{i=1}^{n} p_i^m(n)$, $m \ge 2$, and $\sigma^2 = V(S_n) = \lambda - \lambda_2$. Likewise, let T_n denote a Poisson random variable with expectation λ.

In this section, we study the approximations of the suitably normalized distribution functions F_{S_n} of S_n and F_{T_n} of T_n by a standard normal distribution function Φ. We shall denote by $\varphi = \Phi'$ the standard normal density. Our results will be based on the two following lemmas.

LEMMA 4.1. *There exists an absolute constant A such that, for all $k = 0, 1, \ldots$ and $\lambda > 0$,*

$$\left| F_{T_n}(k) - \Phi(y_k) - \frac{1 - y_k}{6\sqrt{\lambda}} \varphi(y_k) \right| < A\lambda^{-1}, \tag{4.1}$$

where $y_k = \lambda^{-1/2}(k - \lambda + 1/2)$, $k = 0, 1, \ldots$.

PROOF. See Cheng (1964).

LEMMA 4.2. *There exists an absolute constant B such that, for all $n = 1, 2, \ldots$ and $k = 0, 1, \ldots$ we have*

$$\left| F_{S_n}(k) - \Phi(x_k) - \frac{1 - x_k^2}{6\sigma} \varphi(x_k) \left\{ 1 - \frac{2}{\sigma^2}(\lambda_2 - \lambda_3) \right\} \right| \leq B\sigma^{-2}, \tag{4.2}$$

where $x_k = \sigma^{-1/2}(k - \lambda + 1/2)$ and $\sigma = \sqrt{\lambda - \lambda_2}$.

PROOF. See Deheuvels *et al.* (1986).

The following theorem is a consequence of Lemmas 4.1 and 4.2.

THEOREM 4.1. *Let $a = 0.2784\ldots$ be the solution of the equation*

$$1 + x + \log x = 0, \quad a > 0. \tag{4.3}$$

Then if $\lambda \to \infty$ and $\lambda_2/\lambda \to 0$, we have

$$\inf_{\mu, s} d_0(\mathscr{L}(S_n), \mathscr{N}(\mu, s^2)) \sim d\left(\mathscr{L}(S_n), \mathscr{N}\left(\lambda + \frac{2 - a}{3}, \sigma^2 \right) \right) \sim \frac{a}{3\sigma}, \tag{4.4}$$

$$\inf_{\mu, s} d_0(\mathscr{L}(T_n), \mathscr{N}(\mu, s^2)) \sim d\left(\mathscr{L}(T_n), \mathscr{N}\left(\lambda + \frac{2 - a}{3}, \lambda \right) \right) \sim \frac{a}{3\sqrt{\lambda}}, \tag{4.5}$$

where $\mathcal{N}(\mu, s^2)$ *denotes a normal distribution with expectation* μ *and variance* s^2.

PROOF. We prove only the second statement, since the first one may be proved by similar arguments using the observations that $\sigma \sim \sqrt{\lambda}$ and $0 \le \lambda_3 \le \lambda_2 = o(\lambda)$ as $\lambda \to \infty$.

By Taylor's expansion, we see that, for any fixed $M > 0$, we have uniformly over $\max(|\zeta|, |\theta|) \le M$

$$\sup_y \left| \Phi\left(\left(1 + \frac{\zeta}{6\sqrt{\lambda}}\right) y + \frac{\theta}{6\sqrt{\lambda}}\right) - \Phi(y) - \frac{\theta + \zeta y}{6\sqrt{\lambda}} \varphi(y) \right| = O\left(\frac{1}{\lambda}\right).$$

Clearly, the minimum over θ and ζ of $\sup\{|1 - \theta - \zeta y - y^2|\varphi(y)\}$ is reached for $\zeta = 0$ and $1 - \theta = 2a$, where a is the solution of (4.3). This, jointly with (4.1), suffices for our needs. It is noteworthy that $(2 - a)/3 = 0.5738\ldots$ so that the approximation of $\mathcal{L}(S_n)$ by $\mathcal{N}(\lambda + 0.5, \sigma^2)$ is close to optimality in the asymptotic sense given in (4.4).

By (2.10) and (3.16), we see on the other hand that, in the range where $\lambda \to \infty$ and $\lambda_2/\lambda \to 0$, we have $d_0(\mathcal{L}(S_n), \mathcal{L}(T_n)) \sim (1/2\sqrt{2\pi e})(\lambda_2/\lambda)$. A comparison of (4.4) and (4.5) shows that the Poisson approximation of S_n gives an asymptotically better fit (by d_0) than a normal approximation if and only if $\lambda_2 < (2/3)a\sqrt{2\pi e \lambda}$. A precise statement of this result is as follows.

COROLLARY 4.1. *Assume that* $\lambda \to \infty$, $\lambda_2/\lambda \to 0$ *and* $\lambda_2/\lambda^{1/2} \to \alpha$, *where* $0 \le \alpha \le \infty$. *Then,*

(i) *If* $0 \le \alpha < (2/3)a\sqrt{2\pi e} = 0.7670$, *we have ultimately*

$$\inf_{\mu,s} d_0(\mathcal{L}(S_n), \mathcal{N}(\mu, s^2)) > d_0(\mathcal{L}(S_n), \mathcal{L}(T_n)). \tag{4.6}$$

(ii) *If* $(2/3)a\sqrt{2\pi e} < \alpha \le \infty$, *we have ultimately*

$$\inf_{\mu,s} d_0(\mathcal{L}(S_n), \mathcal{N}(\mu, s^2)) < d_0(\mathcal{L}(S_n), \mathcal{L}(T_n)). \tag{4.7}$$

By Corollary 4.1 and the results stated in Section 1, we see that the main range of interest of the Poisson approximation corresponds to the situation where

$$\sum_{i=1}^n p_i^2(n) = O\left\{\left(\sum_{i=1}^n p_i(n)\right)^{1/2}\right\}. \tag{4.8}$$

Since $\lambda_2 \le \lambda$, it is noteworthy that (4.8) is always satisfied when $\lambda = O(1)$.

Acknowledgement

The second author is indebted to the Deutsche Forschungsgemeinschaft for financial support by a Heisenberg Research Grant.

References

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

Barbour, A. D. and Hall, P. (1984). On the rate of Poisson convergence, *Math. Proc. Cambridge Philos. Soc.*, **95**, 473–480.

Cheng, B. (1964). The normal approximation to the Poisson distribution and a proof of a conjecture of Ramanujan, *Bull. Amer. Math. Soc.*, **55**, 396–401.

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

Deheuvels, P. and Pfeifer, D. (1986*b*). Semigroups and Poisson approximation, *Perspectives and New Directions in Theoretical and Applied Statistics*, (eds. M. L. Puri, J. P. Villaplana and W. Wertz), Wiley, New York.

Deheuvels, P. and Pfeifer D. (1987). Operator semigroups and Poisson convergence in selected metrics, *Semigroup Forum*, **34**, 203–224.

Deheuvels, P., Puri, M. L. and Ralescu, S. S. (1986). Asymptotic expansions for sums of non identically distributed binomial random variables, Tech. Report, Laboratoire de Statistique Théorique et Appliquée, Université Paris VI.

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

Pfeifer, D. (1983). A semi-group theoretic proof of Poisson's limit law, *Semigroup Forum*, **26**, 379–383.

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

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

Shorgin, S. Ya. (1977). Approximation of a generalized binomial distribution, *Theory Probab. Appl.*, **22**, 846–850.

Uspensky, J. V. (1931). On Ch. Jordan's series for probability, *Ann. of Math.*, **32**(2), 306–312.

Zolotarev, V. M. (1984). Probability metrics, *Theory Probab. Appl.*, **28**, 278–302.