Hindawi Publishing Corporation
Journal of Inequalities and Applications
Volume 2009, Article ID 965712, 13 pages
doi:10.1155/2009/965712

Research Article

The Kolmogorov Distance between the Binomial and Poisson Laws: Efficient Algorithms and Sharp Estimates

José A. Adell, José M. Anoz, and Alberto Lekuona

Departamento de Métodos Estadísticos, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain

Correspondence should be addressed to José A. Adell, adell@unizar.es

Received 21 May 2009; Accepted 9 October 2009

Recommended by Andrei Volodin

We give efficient algorithms, as well as sharp estimates, to compute the Kolmogorov distance between the binomial and Poisson laws with the same mean λ. Such a distance is eventually attained at the integer part of $\lambda + 1/2 - \sqrt{\lambda + 1/4}$. The exact Kolmogorov distance for $\lambda \le 2 - \sqrt{2}$ is also provided. The preceding results are obtained as a concrete application of a general method involving a differential calculus for linear operators represented by stochastic processes.

Copyright © 2009 José A. Adell et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction and Main Results

There is a huge amount of literature on estimates of different probability metrics between random variables, measuring the rates of convergence in various limit theorems, such as Poisson approximation and the central limit theorem. However, as far as we know, there are only a few papers devoted to obtain exact values for such probability metrics, even in the most simple and paradigmatic examples. In this regard, we mention the results by Kennedy and Quine [1] giving the exact total variation distance between binomial and Poisson distributions, when their common mean λ is smaller than $2 + \sqrt{2}$, approximately, as well as the efficient algorihm provided in the work of Adell et al. [2] to compute this distance for arbitrary values of λ. On the other hand, closed-form expressions for the Kolmogorov and total variation distances between some familiar discrete distributions with different parameters can be found in Adell and Jodrá [3]. Finally, Hipp and Mattner [4] have recently computed the exact Kolmogorov distance in the central limit theorem for symmetric binomial distributions.

The aim of this paper is to obtain efficient algorithms and sharp estimates in the highly classical problem of evaluating the Kolmogorov distance between binomial and Poisson laws having the same mean. The techniques used here are analogous to those in [2] dealing with the total variation distance between the aforementioned laws.

To state our main results, let us introduce some notation. Denote by $\mathbb{Z}_+$ the set of nonnegative integers, $\mathbb{N} = \mathbb{Z}_+ \setminus \{0\}$ and $\mathbb{Z}_n = \{0,1,\dots,n\}$, $n \in \mathbb{N}$. If A is a set of real numbers, 1_A stands for the indicator function of A. For any $x \geq 0$, we set $\lfloor x \rfloor = \max\{k \in \mathbb{Z}_+ : k \leq x\}$ and $\lceil x \rceil = \min\{k \in \mathbb{Z}_+ : x \leq k\}$. For any $m \in \mathbb{Z}_+$, the mth forward differences of a function $\phi : \mathbb{Z}_+ \to \mathbb{R}$ are recursively defined by $\Delta^0\phi = \phi$, $\Delta^1\phi(i) = \phi(i+1) - \phi(i)$, $i \in \mathbb{Z}_+$, and $\Delta^{m+1}\phi = \Delta^1(\Delta^m\phi)$.

Throughout this note, it will be assumed that $n \in \mathbb{N}$, $0 < \lambda < n$, and $p = \lambda/n$. Let $(U_k, k \in \mathbb{N})$ be a sequence of independent identically distributed random variables having the uniform distribution on $[0,1]$. The random variable

$$S_n(t) = \sum_{k=1}^{n} 1_{[0,t]}(U_k), \quad 0 \leq t \leq 1, \; (S_0(t) \equiv 0), \tag{1.1}$$

has the binomial distribution with parameters n and t. Let N_λ be a random variable having the Poisson distribution with mean λ. Recall that the Kolmogorov distance between $S_n(p)$ and N_λ is defined by

$$d(S_n(p), N_\lambda) = \sup_{i \in \mathbb{Z}_+} |f(i)|, \quad f(i) = P(N_\lambda \geq i) - P(S_n(p) \geq i), \; i \in \mathbb{Z}_+. \tag{1.2}$$

Observe that for any $i \in \mathbb{Z}_n$ we have

$$\Delta^1 f(i) = P(S_n(p) = i) - P(N_\lambda = i) = P(N_\lambda = i)\left(\frac{c(n,\lambda)}{g_{n,\lambda}(i)} - 1\right), \tag{1.3}$$

where

$$c(n,\lambda) = n! e^{\lambda}\left(1 - \frac{\lambda}{n}\right)^n, \qquad g_{n,\lambda}(i) = (n-i)!(n-\lambda)^i. \tag{1.4}$$

An efficient algorithm to compute $d(S_n(p), N_\lambda)$ is based on the zeroes of the second Krawtchouk and Charlier polynomials, which are the orthogonal polynomials with respect to the binomial and Poisson distributions, respectively. Interesting references for general orthogonal polynomials are the monographs by Chihara [5] and Schoutens [6].

More precisely, let $k \in \mathbb{N}$ with $k \geq n$, and $0 < t < 1$. The second Krawtchouk polynomial with respect to $S_{k+1}(t)$ is given by

$$Q_2^{(k+1)}(t; x) = \frac{x^2 - (1+2kt)x + k(k+1)t^2}{k(k+1)t^2(1-t)^2}. \tag{1.5}$$

The two zeroes of this polynomial are

$$x_j^{(k+1)}(t) = \frac{1}{2} + kt + (-1)^j\sqrt{kt(1-t) + \frac{1}{4}}, \quad j = 1, 2. \tag{1.6}$$

As $k \to \infty$, $t \to 0$, and $kt \to \lambda$, $Q_2^{(k+1)}(t; x)$ converges to the second Charlier polynomial with respect to N_λ defined by

$$C_2(\lambda; x) = \frac{x^2 - (1+2\lambda)x + \lambda^2}{\lambda^2}, \tag{1.7}$$

the two zeroes of which are

$$r_j(\lambda) = \frac{1}{2} + \lambda + (-1)^j\sqrt{\lambda + \frac{1}{4}}, \quad j = 1, 2. \tag{1.8}$$

Finally, we denote by

$$r_{1,k}(\lambda) = x_1^{(k+1)}\left(\frac{\lambda}{k}\right) = \frac{1}{2} + \lambda - \sqrt{\lambda\left(1 - \frac{\lambda}{k}\right) + \frac{1}{4}}, \tag{1.9}$$

and by

$$r_{2,k}(\lambda) = x_2^{(k+1)}\left(\frac{\lambda}{k+1}\right) = \frac{1}{2} + \lambda\frac{k}{k+1} + \sqrt{\lambda\left(1 - \frac{\lambda}{k+1}\right)\frac{k}{k+1} + \frac{1}{4}} \tag{1.10}$$

the smallest zero of $Q_2^{(k+1)}(\lambda/k; x)$ and the greatest zero of $Q_2^{(k+1)}(\lambda/(k+1); x)$, respectively, (see Figure 1).

Our first main result is the following.

Theorem 1.1. *Let $n \in \mathbb{N}$ and $0 < \lambda < n$. Then,*

$$d(S_n(p), N_\lambda) = \max\{-f(l_\lambda(n)), f(m_\lambda(n) + 1)\}, \tag{1.11}$$

where f is defined in (1.2),

$$l_\lambda(n) = \min\{i \in [\lfloor r_1(\lambda)\rfloor + 1, \lceil r_{1,n}(\lambda)\rceil] \cap \mathbb{Z}_n; g_{n,\lambda}(i) \le c(n, \lambda)\}, \tag{1.12}$$

$$m_\lambda(n) = \max\{i \in [\lfloor r_{2,n}(\lambda)\rfloor, \lceil r_2(\lambda)\rceil - 1] \cap \mathbb{Z}_n; g_{n,\lambda}(i) \le c(n, \lambda)\}. \tag{1.13}$$

Looking at Figure 1 and taking into account (1.8), (1.9), and (1.12) we see the following. The number of computations needed to evaluate $l_\lambda(n)$ is approximately $r_{1,n}(\lambda) - r_1(\lambda)$, that is, $\lambda\sqrt{\lambda}/(2n)$, approximately. This last quantity is relatively small, since N_λ

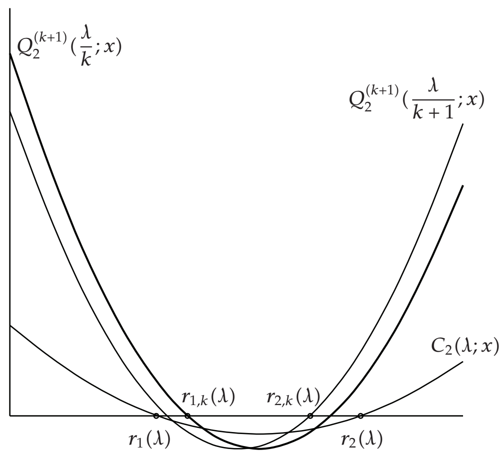

Figure 1: The polynomials $Q_2^{(k+1)}(\lambda/k;x)$, $Q_2^{(k+1)}(\lambda/(k+1);x)$, and $C_2(\lambda;x)$, for $\lambda > 2$.

approximates $S_n(p)$ if and only if $p = \lambda/n$ is close to zero. Moreover, the set $[\lfloor r_1(\lambda)\rfloor + 1, \lceil r_{1,n}(\lambda)\rceil] \cap \mathbb{Z}_n$ has two points at most, whenever $r_{1,n}(\lambda) - r_1(\lambda) < 1$, and this happens if

$$n > \frac{\lambda^2}{\sqrt{4\lambda+5}-2}. \tag{1.14}$$

As follows from (1.2), the natural way to compute the Kolmogorov distance $d(S_n(p), N_\lambda)$ is to look at the maximum absolute value of the function

$$f(i) = \sum_{k=0}^{i-1} \Delta^1 f(k) = \sum_{k=0}^{i-1} \big(P\big(S_n(p) = k\big) - P(N_\lambda = k)\big), \quad i \in \mathbb{N}. \tag{1.15}$$

From a computational point of view, the main question is to ask how many evaluations of the probability differences $P(S_n(p) = k) - P(N_\lambda = k)$ are required to exactly compute $d(S_n(p), N_\lambda)$. According to Theorem 1.1 and (1.8), the number of such evaluations is $\lambda - \sqrt{\lambda}$ at least, and $\lambda + \sqrt{\lambda}$ at most, approximately.

On the other hand, $r_{1,n}(\lambda)$ and $r_{2,n}(\lambda)$ converge, respectively, to $r_1(\lambda)$ and $r_2(\lambda)$, as $n \to \infty$. Thus, Theorem 1.1 leads us to the following asymptotic result.

Corollary 1.2. *Let $n \in \mathbb{N}$ and $0 < \lambda < n$. Let $n_0(\lambda)$ be the smallest integer such that $\lceil r_{1,n}(\lambda)\rceil = \lfloor r_1(\lambda)\rfloor + 1$ and $\lfloor r_{2,n}(\lambda)\rfloor = \lceil r_2(\lambda)\rceil - 1$, for $n \ge n_0(\lambda)$. Then, one has for any $n \ge n_0(\lambda)$*

$$\begin{aligned} d\big(S_n(p), N_\lambda\big) = \max\big\{&P(N_\lambda \le \lfloor r_1(\lambda)\rfloor) - P\big(S_n(p) \le \lfloor r_1(\lambda)\rfloor\big), \\ &P\big(S_n(p) \le \lceil r_2(\lambda)\rceil\big) - P(N_\lambda \le \lceil r_2(\lambda)\rceil)\big\}. \end{aligned} \tag{1.16}$$

Unfortunately, $n_0(\lambda)$ is not uniformly bounded when λ varies in an arbitrary compact set. In fact, since $r_1(l + \sqrt{l}) = l$, $l \in \mathbb{N}$, and $r_2(m - \sqrt{m}) = m$, $m = 2, 3, \ldots$, it can be verified that $n_0(\lambda) \to \infty$, when $\lambda \to l + \sqrt{l}$ from the left, $l \in \mathbb{N}$, or when $\lambda \to m - \sqrt{m}$ from the right, $m = 2, 3, \ldots$. This explains why $l_\lambda(n)$ and $m_\lambda(n)$ in Theorem 1.1 have no simple form in general.

Finally, it may be of interest to compare Theorem 1.1 and Corollary 1.2 with the exact value of the Kolmogorov distance in the central limit theorem for symmetric binomial distributions obtained by Hipp and Mattner [4]. These authors have shown that (cf. [4, Corollary 1.1])

$$d\left(\frac{S_n(1/2) - n/2}{\sqrt{n/4}}, Z\right) = \begin{cases} P\left(Z \leq \frac{1}{\sqrt{n}}\right) - \frac{1}{2} & (n \text{ odd}), \\ \frac{1}{2} P\left(S_n\left(\frac{1}{2}\right) = \frac{n}{2}\right) & (n \text{ even}), \end{cases} \tag{1.17}$$

where Z is a standard normal random variable. Roughly speaking, (1.17) tells us that the Kolmogorov distance in this version of the central limit theorem is attained at the mean of the respective distributions; whereas according to Theorem 1.1 and Corollary 1.2, the Kolmogorov distance in our Poisson approximation setting is attained at the mean ± the standard deviation of the corresponding distributions.

For small values of λ, we are able to give the following closed-form expression.

Corollary 1.3. *Let $n \in \mathbb{N}$. If $0 < \lambda \leq 2 - \sqrt{2}$, then*

$$d(S_n(p), N_\lambda) = P(N_\lambda = 0) - P(S_n(p) = 0) = e^{-\lambda} - \left(1 - \frac{\lambda}{n}\right)^n. \tag{1.18}$$

Corollary 1.3 can be seen as a counterpart of the total variation result established by Kennedy and Quine [1, Theorem 1], stating that

$$d_{\mathrm{TV}}(S_n(p), N_\lambda) = P(S_n(p) = 1) - P(N_\lambda = 1) = \lambda\left(\left(1 - \frac{\lambda}{n}\right)^{n-1} - e^{-\lambda}\right), \tag{1.19}$$

for any $n \in \mathbb{N}$ and $0 < \lambda \leq 2 - \sqrt{2}$, where $d_{\mathrm{TV}}(\cdot, \cdot)$ stands for the total variation distance.

For any $m \in \mathbb{N}$, $n = 2, 3, \ldots$, and $0 < \lambda < n$, we denote by

$$K_\lambda(n) = \frac{n+2}{2(n+1)}\left(\frac{2\lambda}{3} f_3(n, \lambda) + \frac{\lambda^2}{4} f_4(n, \lambda)\right), \tag{1.20}$$

where

$$f_m(n, \lambda) = \min\left(2^{m-1}, \frac{1}{2}\left(\frac{n+2}{n-1}\right)^{3/2} \sqrt{\frac{m!}{\lambda^m (1 - (\lambda/n))^m}}\right). \tag{1.21}$$

Sharp estimates for the Kolmogorov distance are given in the following.

Theorem 1.4. *Let $n = 2, 3, \ldots, 0 < \lambda < n$ and $p = \lambda/n$. Then,*

$$\left| d(S_n(p), N_\lambda) - \frac{1}{2} p M_\lambda \right| \le K_\lambda(n) p^2, \tag{1.22}$$

where

$$M_\lambda = e^{-\lambda} \frac{\lambda^{\lfloor r_1(\lambda) \rfloor} (\lambda - \lfloor r_1(\lambda) \rfloor)}{\lfloor r_1(\lambda) \rfloor !}. \tag{1.23}$$

Upper bounds for the Kolmogorov distance in Poisson approximation for sums of independent random indicators have been obtained by many authors using different techniques. We mention the following estimates in the case at hand:

$$d(S_n(p), N_\lambda) \le \frac{1}{2} \lambda p \tag{1.24}$$

(Serfling [7]),

$$d(S_n(p), N_\lambda) \le \frac{\pi}{4} p \tag{1.25}$$

(Hipp [8]),

$$\left| d(S_n(p), N_\lambda) - \frac{1}{2} p \max\left\{ M_\lambda, \widetilde{M}_\lambda \right\} \right| \le \left(\frac{1}{2} + \sqrt{\frac{\pi}{8}} \right) \frac{p^{3/2}}{1 - \sqrt{p}} \tag{1.26}$$

(Deheuvels et al. [9]),

$$d(S_n(p), N_\lambda) \le \frac{1}{2e} p + \frac{6p^{3/2}}{5(1 - \sqrt{p})} \tag{1.27}$$

(Roos [10]), where

$$\widetilde{M}_\lambda = e^{-\lambda} \frac{\lambda^{\lfloor r_2(\lambda) \rfloor} (\lfloor r_2(\lambda) \rfloor - \lambda)}{\lfloor r_2(\lambda) \rfloor !}, \tag{1.28}$$

and the constant $1/(2e)$ in the last estimate is best possible (cf. Roos [10]). It is readily seen from (1.23) that

$$\lim_{\lambda \to \infty} M_\lambda = \frac{1}{\sqrt{2\pi e}}, \quad M_\lambda = \lambda e^{-\lambda}, \ 0 < \lambda < 2. \tag{1.29}$$

On the other hand, it follows from Roos [10] and Lemma 2.1 below that

$$\widetilde{M}_\lambda \le M_\lambda \le M_1 = \frac{1}{e}, \quad \lambda > 0. \tag{1.30}$$

Table 1: Upper bounds for $d(S_n(p), N_\lambda)$: Serfling (S), Hipp (H), Deheuvels et al. (D), Roos (R), and Adell et al. (A).

n	p	S	H	D	R	A
100		0.0050	0.007854	0.003091	0.003173	**0.001916**
200	0.01	0.0100	0.007854	0.002605	0.003173	**0.001416**
500		0.0250	0.007854	0.002656	0.003173	**0.001454**
1000		0.0500	0.007854	0.002603	0.003173	**0.001396**
200		0.0025	0.003927	0.001348	0.001376	**0.000938**
400	0.005	0.0050	0.003927	0.001105	0.001376	**0.000692**
1000		0.0125	0.003927	0.001131	0.001376	**0.000714**
2000		0.0250	0.003927	0.001104	0.001376	**0.000687**

Such properties, together with simple numerical computations performed with Maple™ 9.01, show that estimate (1.22) is always better than the preceding ones for $0 < p \leq 1/3$ and $n \geq 10$. Numerical comparisons are exhibited in Table 1.

On the other hand, the referee has drawn our attention to a recent paper by Vaggelatou [11], where the author obtains upper bounds for the Kolmogorov distance between sums of independent integer-valued random variables. Specializing Corollary 15 in [11] to the case at hand, Vaggelatou gives the upper bound

$$d(S_n(p), N_\lambda) \leq \frac{M_\lambda}{2(1-2(1-e^{-p}))}p + \frac{\lambda^2}{2}p^2. \tag{1.31}$$

Comparing Corollary 1.3 and (1.22) with (1.31), we see the following. The constant in the main term of the order of p in (1.22) is better than that in (1.31). The constant $K_\lambda(n)$ in the remainder term of the order of p^2 in (1.22) is uniformly bounded in λ, whereas; $\lambda^2/2$ is not. However, $\lambda^2/2$ is better than $K_\lambda(n)$ for small values of $\lambda > 2-\sqrt{2}$ (recall that Corollary 1.3 gives the exact distance for $0 < \lambda \leq 2-\sqrt{2}$). As a result, for moderate or large values of n, estimate (1.31) is sometimes better than (1.22) for $2-\sqrt{2} < \lambda < 1$, approximately. Otherwise, Corollary 1.3 and (1.22) provide better bounds than (1.31). This is illustrated in Table 2.

We finally establish that, for small values of p, the Kolmogorov distance is attained at $\lfloor r_1(\lambda) \rfloor$, that is, at $\lambda - \sqrt{\lambda}$, approximately. This completes the statement in Corollary 1.3.

Corollary 1.5. *For any $\lambda > 0$, one has*

$$\lim_{p\to 0}\frac{1}{p}d(S_n(p), N_\lambda) = \lim_{p\to 0}\frac{1}{p}\left(P(N_\lambda \leq \lfloor r_1(\lambda)\rfloor) - P(S_n(p) \leq \lfloor r_1(\lambda)\rfloor)\right) = \frac{M_\lambda}{2}. \tag{1.32}$$

Remark 1.6. As far as upper bounds are concerned, the methods used in this paper can be adapted to cover more general cases referring to Poisson approximation (see, e.g., the Introduction in [2] and the references therein). However, the obtention of efficient algorithms leading to exact values is a more delicate question. As we will see in Section 2, specially in formula (2.1), such a problem is based on two main facts: first, the explicit form of the orthogonal polynomials associated to the random variables to be approximated, and, second, the relation between expectations involving forward differences and expectations involving

Table 2: Upper bounds for $d(S_n(p), N_\lambda)$: Vaggelatou (V) and Adell et al. (A).

λ	n	V	A
0.6	20	**0.0054116**	0.0059268
	50	**0.0020499**	0.0021122
	100	**0.0010063**	0.0010199
	200	**0.0004985**	0.0005017
	500	**0.0001983**	0.0001988
	1000	**0.0000990**	0.0000991
0.9	20	**0.0098476**	0.0103392
	50	**0.0035463**	0.0035676
	100	**0.0017095**	0.0017106
	200	0.0008390	**0.0008388**
	500	0.0003318	**0.0003317**
	1000	0.0001653	**0.0001653**
1	20	**0.0114410**	0.0117428
	50	0.0040305	**0.0040086**
	100	0.0019267	**0.0019162**
	200	0.0009415	**0.0009382**
	500	0.0003714	**0.0003708**
	1000	0.0001848	**0.0001847**
2	20	0.0367148	**0.0228808**
	50	0.0090741	**0.0065533**
	100	0.0036183	**0.0029671**
	200	0.0015808	**0.0014156**
	500	0.0005777	**0.0005510**
	1000	0.0002798	**0.0002731**

these orthogonal polynomials. For instance, an explicit expression for the orthogonal polynomials associated to general sums of independent random indicators seems to be unknown.

2. The Proofs

The key tool to prove the previous results is the following formula established in [2, formula (1.4)]. For any function $\phi : \mathbb{Z}_+ \to \mathbb{R}$ for which the expectations below exist, we have

$$\begin{aligned} E\phi(S_n(p)) - E\phi(N_\lambda) &= -\lambda^2 \sum_{k=n}^{\infty} \frac{1}{k(k+1)} EU\Delta^2\phi(S_{k-1}(T_k)) \\ &= -\lambda^2 \sum_{k=n}^{\infty} \frac{1}{k(k+1)} EU\phi(S_{k+1}(T_k))Q_2^{(k+1)}(T_k; S_{k+1}(T_k)), \end{aligned} \tag{2.1}$$

where

$$T_k = \frac{\lambda}{k}\left(1 - \frac{UV}{k+1}\right), \quad k = n,\ n+1, \ldots, \tag{2.2}$$

and U and V are independent identically distributed random variables having the uniform distribution on $[0,1]$, also independent of the sequence $(U_k, k \in \mathbb{N})$ in (1.1).

Proof of Theorem 1.1. Let $n \in \mathbb{N}$, $0 < \lambda < n$, and $i \in \mathbb{Z}_+$. The function $g_{n,\lambda}(i)$ defined in (1.4) decreases in $[0, \lfloor \lambda + 1 \rfloor] \cap \mathbb{Z}_n$ and increases in $[\lfloor \lambda + 1 \rfloor, n] \cap \mathbb{Z}_n$. This property, together with definitions (1.2)–(1.4), readily implies the following. There are integers $1 \le l_\lambda(n) \le m_\lambda(n) \le n$ such that

$$\{i \in \mathbb{Z}_n; g_{n,\lambda}(i) \le c(n,\lambda)\} = \left\{i \in \mathbb{Z}_+ : \Delta^1 f(i) \ge 0\right\} = [l_\lambda(n), m_\lambda(n)] \cap \mathbb{Z}_n. \tag{2.3}$$

As a consequence of (2.3), the function $f(i)$ defined in (1.2) starts from $f(0) = 0$, decreases in $[0, l_\lambda(n)]$, increases in $[l_\lambda(n), m_\lambda(n) + 1]$, decreases in $[m_\lambda(n) + 1, \infty)$, and tends to zero as $i \to \infty$. We therefore conclude that

$$d(S_n(p), N_\lambda) = \max\{-f(l_\lambda(n)), f(m_\lambda(n) + 1)\}. \tag{2.4}$$

To show (1.12) and (1.13), we apply the second equality in (2.1) to the function $\phi = 1_{\{i\}}$, thus obtaining by virtue of (1.3)

$$\Delta^1 f(i) = -\lambda^2 \sum_{k=n}^{\infty} \frac{1}{k(k+1)} EU 1_{\{i\}}(S_{k+1}(T_k)) Q_2^{(k+1)}(T_k; i). \tag{2.5}$$

In view of (2.3), statements (1.12) and (1.13) will follow as soon as we show that

$$\Delta^1 f(i) \ge 0, \quad \text{if } i \in I_\lambda(n) = [\lceil r_{1,n}(\lambda) \rceil, \lfloor r_{2,n}(\lambda) \rfloor] \cap \mathbb{Z}_n, \tag{2.6}$$

as well as

$$\Delta^1 f(i) < 0, \quad \text{if } i \in J_\lambda(n) = ([0, \lfloor r_1(\lambda) \rfloor] \cup [\lceil r_2(\lambda) \rceil, \infty)) \cap \mathbb{Z}_n. \tag{2.7}$$

Observe that some of the sets in (2.6) and (2.7) could be empty. To this end, let $k \in \mathbb{N}$ with $k \ge n$, and $\lambda/(k+1) \le t \le \lambda/k$. Since the functions $x_j^{(k+1)}(t)$ defined in (1.6) are increasing in t, we have by virtue of (1.9) and (1.10)

$$\begin{aligned} x_1^{(k+1)}(t) &\le x_1^{(k+1)}\left(\frac{\lambda}{k}\right) = r_{1,k}(\lambda) \le r_{1,n}(\lambda) \le r_{2,n}(\lambda) \\ &\le r_{2,k}(\lambda) = x_2^{(k+1)}\left(\frac{\lambda}{k+1}\right) \le x_2^{(k+1)}(t). \end{aligned} \tag{2.8}$$

Again by (1.9) and (1.10), this means that $Q_2^{(k+1)}(t; i) \le 0$, for any $i \in I_\lambda(n)$. This fact, in conjunction with (2.2) and (2.5), shows (2.6).

To prove (2.7), we distinguish the following two cases.

Case 1 ($\lambda > 2$). By (1.6), (1.8), (1.9), and (1.10), we have

$$r_1(\lambda) < x_1^{(k+1)}\left(\frac{\lambda}{k+1}\right) \leq x_1^{(k+1)}(t) < x_2^{(k+1)}(t) \leq x_2^{(k+1)}\left(\frac{\lambda}{k}\right) < r_2(\lambda), \tag{2.9}$$

which implies that $Q_2^{(k+1)}(t;i) > 0$, for any $i \in J_n(\lambda)$. As before, this property shows (2.7).

Case 2 ($\lambda \leq 2$). In this occasion, we have $x_1^{(k+1)}(\lambda/(k+1)) \leq r_1(\lambda) \leq 1$. Since $\Delta^1 f(0) < 0$ and the remaining inequalities in (2.9) are satisfied, we conclude as in the previous case that (2.7) holds. The proof is complete. □

Proof of Corollary 1.3. For $0 < \lambda \leq 2 - \sqrt{2}$, (1.8) implies that $\lfloor r_1(\lambda) \rfloor + 1 = \lceil r_2(\lambda) \rceil - 1 = 1$, and, therefore, $l_\lambda(n) = m_\lambda(n) = 1$, as follows from Theorem 1.1. By (1.11), this in turn implies that

$$d(S_n(p), N_\lambda) = \max\{-f(1), f(2)\}. \tag{2.10}$$

On the other hand, we have from (1.2)

$$-f(1) - f(2) = E\psi(N_\lambda) - E\psi(S_n(p)), \tag{2.11}$$

where $\psi : \mathbb{Z}_+ \to \mathbb{R}$ is the convex function given by $\psi(i) = 2 \cdot 1_{\{0\}}(i) + 1_{\{1\}}(i)$, $i \in \mathbb{Z}_+$. Since $\Delta^2\psi \geq 0$, the first inequality in (2.1) proves that the right-hand side in (2.11) is nonnegative. This, together with (2.10), shows that $d(S_n(p), N_\lambda) = -f(1)$ and completes the proof. □

Let $n = 2, 3, \ldots,\ 0 < \lambda < n$, and $p = \lambda/n$. For any function $\phi : \mathbb{Z}_+ \to [0,1]$, we have

$$E\Delta^2\phi(N_\lambda) = E\phi(N_\lambda)C_2(\lambda; N_\lambda), \tag{2.12}$$

$$\left| E\phi(S_n(p)) - E\phi(N_\lambda) + \frac{p}{2}\lambda E\phi(N_\lambda)C_2(\lambda; N_\lambda) \right| \leq K_\lambda(n)p^2, \tag{2.13}$$

where $K_\lambda(n)$ is defined in (1.20). Formula (2.12) can be found in Barbour et al. [12, Lemma 9.4.4]; whereas estimate (2.13) is established in Adell et al. [2, formula (6.1)]. Choosing $\phi = 1_{[i,\infty)},\ i \in \mathbb{Z}_+$ in (2.12), we consider the function

$$\begin{aligned} g(i) &= \lambda E 1_{[i,\infty)}(N_\lambda)C_2(\lambda; N_\lambda) \\ &= \lambda\left(E1_{\{i-2\}}(N_\lambda) - E1_{\{i-1\}}(N_\lambda)\right), \quad i \in \mathbb{Z}_+. \end{aligned} \tag{2.14}$$

Observe that

$$\Delta^1 g(i) = -\lambda 1_{\{i\}}(N_\lambda)C_2(\lambda; N_\lambda), \quad i \in \mathbb{Z}_+. \tag{2.15}$$

Therefore, the function $g(\cdot)$ in (2.14) starts from $g(0) = 0$, decreases in $[0, \lfloor r_1(\lambda)\rfloor + 1]$, increases in $[\lfloor r_1(\lambda) + 1\rfloor, \lfloor r_2(\lambda)\rfloor + 1]$, and decreases to zero in $[\lfloor r_2(\lambda)\rfloor + 1, \infty)$. We therefore have from (2.14)

$$\sup_{i\in\mathbb{Z}_+}|g(i)| = \max\{-g(\lfloor r_1(\lambda)\rfloor + 1), g(\lfloor r_2(\lambda)\rfloor + 1)\} = \max\left\{M_\lambda, \widetilde{M}_\lambda\right\}, \tag{2.16}$$

where M_λ and $\widetilde{M}_\lambda$ are defined in (1.23) and (1.28), respectively.

As shown in the following auxiliary result, it turns out that $M_\lambda \geq \widetilde{M}_\lambda$, $\lambda > 0$. In this respect, we will need the well-known inequalities

$$B_{2n}(x) \leq \log(1+x) \leq B_{2n+1}(x), \qquad B_n(x) = -\sum_{k=1}^{n}\frac{(-x)^k}{k}, \tag{2.17}$$

for $n \in \mathbb{N}$ and $0 < x < 1$.

Lemma 2.1. *For any $\lambda > 0$, one has $M_\lambda \geq \widetilde{M}_\lambda$. In addition, for any $\lambda \geq 2$, one has $M_\lambda > (2\pi e)^{-1/2} > \widetilde{M}_\lambda$.*

Proof. We will only show that $M_\lambda > (2\pi e)^{-1/2}$, $\lambda \geq 2$, with the proof of the remaining inequalities being similar. Let $m \in \mathbb{N}$. Since the function $r_1(\cdot)$ defined in (1.8) is increasing and $r_1(m + \sqrt{m}) = m$, we see that

$$M_\lambda = e^{-\lambda}\frac{\lambda^m(\lambda - m)}{m!}, \quad \lambda \in \left[m + \sqrt{m}, m + 1 + \sqrt{m+1}\right). \tag{2.18}$$

As follows by calculus, in each interval $[m+\sqrt{m}, m+1+\sqrt{m+1})$, M_λ attains its minimum at the endpoints. On the other hand, M_λ converges to $(2\pi e)^{-1/2}$, as $\lambda \to \infty$. Therefore, it will be enough to show that the sequence $(\log M_{m+\sqrt{m}}, m \in \mathbb{N})$ is decreasing, or, in other words, that

$$\begin{aligned}&(m+1)\log\left(1 + \frac{1}{\sqrt{m+1}}\right) - \sqrt{m+1} - \left(m\log\left(1+\frac{1}{\sqrt{m}}\right) - \sqrt{m}\right)\\&\quad + \left(m + \frac{1}{2}\right)\log\left(1 + \frac{1}{m}\right) - 1 < 0.\end{aligned} \tag{2.19}$$

Simple numerical computations show that (2.19) holds for $1 \leq m \leq 6$. Assume that $m \geq 7$. By (2.17), the left-hand side in (2.19) is bounded above by

$$\begin{aligned}&(m+1)B_5\left(\frac{1}{\sqrt{m+1}}\right) - \sqrt{m+1} - \left(mB_6\left(\frac{1}{\sqrt{m}}\right) - \sqrt{m}\right)\\&\quad + \left(m+\frac{1}{2}\right)B_3\left(\frac{1}{m}\right) - 1 = \frac{1}{3}\left(\frac{1}{\sqrt{m+1}} - \frac{1}{\sqrt{m}}\right) - \frac{1}{4}\left(\frac{1}{m+1} - \frac{1}{m}\right)\\&\quad + \frac{1}{5}\left(\frac{1}{(m+1)\sqrt{m+1}} - \frac{1}{m\sqrt{m}}\right) + \frac{1}{4m^2} + \frac{1}{6m^3} < 0.\end{aligned} \tag{2.20}$$

This completes the proof. □

Proof of Theorem 1.4. Applying (2.13) to $\phi = 1_{[i,\infty)}$, $i \in \mathbb{Z}_+$, and using the converse triangular inequality for the usual sup-norm, we obtain

$$\left| d(S_n(p), N_\lambda) - \frac{p}{2} \sup_{i\in\mathbb{Z}_+} |g(i)| \right| \le \sup_{i\in\mathbb{Z}_+} \left| E1_{[i,\infty)}(S_n(p)) - E1_{[i,\infty)}(N_\lambda) + \frac{p}{2} g(i) \right| \le K_\lambda(n) p^2. \tag{2.21}$$

Thus, the conclusion follows from (2.16) and Lemma 2.1. □

We have been aware that Boutsikas and Vaggelatou have recently provided in [13] an independent proof of Lemma 2.1.

Proof of Corollary 1.5. From (2.16) and the orthogonality of $C_2(\lambda;\cdot)$, we get

$$\lambda E1_{[0,\lfloor r_1(\lambda)\rfloor]}(N_\lambda) C_2(\lambda; N_\lambda) = -g(\lfloor r_1(\lambda)\rfloor + 1) = M_\lambda. \tag{2.22}$$

Therefore, applying (2.13) to the function $\phi = -1_{[0,\lfloor r_1(\lambda)\rfloor]}$, as well as Theorem 1.4, we obtain the desired conclusion. □

Acknowledgments

The authors thank the referees for their careful reading of the manuscript and for their remarks and suggestions, which greatly improved the final outcome. This work has been supported by Research Grants MTM2008-06281-C02-01/MTM and DGA E-64, and by FEDER funds.

References

[1] J. E. Kennedy and M. P. Quine, "The total variation distance between the binomial and Poisson distributions," *The Annals of Probability*, vol. 17, no. 1, pp. 396–400, 1989.

[2] J. A. Adell, J. M. Anoz, and A. Lekuona, "Exact values and sharp estimates for the total variation distance between binomial and Poisson distributions," *Advances in Applied Probability*, vol. 40, no. 4, pp. 1033–1047, 2008.

[3] J. A. Adell and P. Jodrá, "Exact Kolmogorov and total variation distances between some familiar discrete distributions," *Journal of Inequalities and Applications*, vol. 2006, Article ID 64307, 8 pages, 2006.

[4] C. Hipp and L. Mattner, "On the normal approximation to symmetric binomial distributions," *Theory of Probability and Its Applications*, vol. 52, no. 3, pp. 516–523, 2008.

[5] T. S. Chihara, *An Introduction to Orthogonal Polynomials*, Gordon and Breach, New York, NY, USA, 1978.

[6] W. Schoutens, *Stochastic Processes and Orthogonal Polynomials*, vol. 146 of *Lecture Notes in Statistics*, Springer, New York, NY, USA, 2000.

[7] R. J. Serfling, "Some elementary results on Poisson approximation in a sequence of Bernoulli trials," *SIAM Review*, vol. 20, no. 3, pp. 567–579, 1978.

[8] C. Hipp, "Approximation of aggregate claims distributions by compound Poisson distributions," *Insurance: Mathematics & Economics*, vol. 4, no. 4, pp. 227–232, 1985.

[9] P. Deheuvels, D. Pfeifer, and M. L. Puri, "A new semigroup technique in Poisson approximation," *Semigroup Forum*, vol. 38, no. 2, pp. 189–201, 1989.

[10] B. Roos, "Sharp constants in the Poisson approximation," *Statistics & Probability Letters*, vol. 52, no. 2, pp. 155–168, 2001.

[11] E. Vaggelatou, “A new method for bounding the distance between sums of independent integer-valued random variables,” *Methodology and Computing in Applied Probability*. In press.

[12] A. D. Barbour, L. Holst, and S. Janson, *Poisson Approximation*, vol. 2 of *Oxford Studies in Probability*, The Clarendon Press, Oxford University Press, New York, NY, USA, 1992.

[13] M. V. Boutsikas and E. Vaggelatou, “A new method for obtaining sharp compound Poisson approximation error estimates for sums of locally dependent random variables,” to appear in *Bernoulli*.