

POISSON APPROXIMATION TO THE CONVOLUTION OF POWER SERIES  
DISTRIBUTIONS

BY

A. N. KUMAR (MUMBAI), P. VELLAISAMY (MUMBAI), AND F. VIENS (EAST  
LANSING)

*Abstract.* In this article, we obtain, for the total variation distance, the error bounds between Poisson and convolution of power series distributions via Stein's method. This provides a unified approach to many known discrete distributions. Several Poisson limit theorems follow as corollaries from our bounds. As applications, we compare the Poisson approximation results with the negative binomial approximation results, for the sums of Bernoulli, geometric, and logarithmic series random variables.

**2010 AMS Mathematics Subject Classification:** Primary: 62E17, 62E20; Secondary: 60F05, 60E05.

**Key words and phrases:** Convolution of distributions, Poisson and negative binomial approximation, power series distribution, Stein's method.

1. INTRODUCTION

The convolution of distributions play an important role in several applications in the areas related to rare events, waiting time, and wireless communications, among many others. It is in general difficult to find the distributions of the sums of especially independent and non-identical random variables (rvs). In such cases, approximations to a known distribution is useful in applications. For example, Poisson approximation to the convolution of Bernoulli rvs is studied by Barbour and

Hall [2], Chen [4], Kerstan [10], and Le Cam [11]. The Poisson approximation to the convolution of geometric rvs is studied by, for example, Barbour [1], Hung and Giang [7] and Teerapabolarn and Wongkasem [15]. The Poisson approximation to the convolution of negative binomial rvs is studied by Teerapabolarn [14], and Vellaisamy and Upadhye [17], among others.

In this article, we focus on the convolution of non-identically distributed power series distributions (PSDs) and obtain the upper bounds for its approximation to Poisson distribution, using Stein's method. The distance metric used is the total variation distance. We establish a device (Theorem 3.1) where, beyond moment matching, the main question is to compare a first-moment-type statistic for the PSDs under consideration, with the same statistic for derived PSDs, where the two statistics would coincide if the laws were all Poisson distributed. We show that the limit theorems given by Pérez-Abreu [13] follow from our results, as special cases. As examples, we discuss the Poisson convergence result for binomial, negative binomial and logarithmic series distributions. Furthermore, we mention the negative binomial approximation results and compare the bounds with Poisson approximation results, either theoretically or numerically. It is shown that our bounds are either comparable to or an improvement over the existing bounds. We have discussed also the Poisson approximation to the convolution of logarithmic series distributions, which has not been studied so far in the literature.

The article is organized as follows. In Section 2, we discuss some known results for PSDs and Stein's method. In Section 3, we derive the error bounds between Poisson and the convolution of PSDs, and discuss some relevant consequences. In Section 4, we present the results for negative binomial approximation to PSDs obtained from Vellaisamy *et al.* [18]. Finally, we give a numerical comparison between Poisson and negative binomial approximation to the PSDs.

## 2. PRELIMINARIES

First we introduce the notation and briefly discuss the Stein's method. Let  $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ , the set of non-negative integers, and  $Z$  be a random variable (rv) with

$$(2.1) \quad \mathbb{P}(Z = k) = \frac{a_k \theta^k}{h(\theta)}, \quad k \in \mathbb{Z}_+, \quad \theta > 0,$$

where  $a_k \geq 0$  and  $h(\theta) = \sum_{k=0}^{\infty} a_k \theta^k$ . Then we say  $Z$  belongs to the class of PSDs corresponding to  $h(\theta)$ . It can be easily verified that the Bernoulli, binomial, geometric, negative binomial and logarithmic series distributions, among many others, belong to the class of PSDs. For more details, we refer the reader to Johnson *et al.* [8]. Through out the paper,  $\text{Poi}(\lambda)$  and  $\text{Geo}(p)$  denote respectively the Poisson and geometric distributions.

Next, we describe briefly Stein's method to obtain the error bound, under the total variation norm, between the distributions of two discrete rvs  $Y$  and  $Z$ , where the former plays the role of the target law, and the estimation in total variation is interpreted as an approximation of the latter law by the target law. Stein's method involves mainly the following three steps:

1. Identify a so-called Stein operator for the target law of  $Y$ . Specifically, define the class of bounded real-valued functions on  $\mathbb{Z}_+$  as  $\mathcal{G} = \{f : \mathbb{Z}_+ \rightarrow \mathbb{R} \mid f \text{ is bounded}\}$ . For a  $\mathbb{Z}_+$ -valued rv  $Y$ , define  $\mathcal{G}_Y = \{g \in \mathcal{G} \mid g(0) = 0 \text{ and } g(x) = 0, \text{ for } x \notin \text{Supp}(Y)\}$ . An operator  $\mathcal{A}_Y$  defined on  $\mathcal{G}_Y$  is called a Stein operator for the rv  $Y$  if

$$(2.2) \quad \mathbb{E}[\mathcal{A}_Y g(Y)] = 0, \quad \text{for } g \in \mathcal{G}_Y.$$

2. Solve the so-called Stein equations. Specifically, for any fixed test function  $f \in \mathcal{G}$  of interest (see item 3 below), and for a fixed target law of  $Y$ , we solve for  $g$  in the following functional equation:

$$(2.3) \quad \mathcal{A}_Y g(k) = f(k) - \mathbb{E}f(Y), \quad f \in \mathcal{G} \text{ and } g \in \mathcal{G}_Y.$$

This equation is called the Stein equation corresponding to  $f$  and to the Stein operator  $\mathcal{A}_Y$ .

3. Replace  $k$  by a rv  $Z$  in Stein equation (2.3), and take expectation and supremum over all test functions  $f$  of interest, to get

$$(2.4) \quad d_{TV}(Z, Y) := \sup_{f \in \mathcal{H}} |\mathbb{E}f(Z) - \mathbb{E}f(Y)| = \sup_{f \in \mathcal{H}} |\mathbb{E}\mathcal{A}_Y g(Z)|,$$

where  $\mathcal{H} = \{I(A) | A \subseteq \mathbb{Z}\}$  and  $I(A)$  is the indicator function of  $A$  and  $d_{TV}(Z, Y)$  is the total variation distance between the distributions of  $Z$  and  $Y$ . Note that in principle,  $g$  depends on  $f$  via the Stein equation (2.3), even if a solution to this equation exists and is not unique. However, any upper bound on the behavior of  $|\mathcal{A}_Y g|$ , which in practice may be global and may not depend on how  $g$  depends on  $f$ , can then be exploited to get a convenient upper bound on the total variation distance of interest. See an instance of this phenomenon below in equations (2.6) and (2.7).

The above methodology also applies to the use of Stein's equations for continuous variables, classically for comparisons with normal laws, but also for any number of laws in specific classes which include the normal. See for instance the research monograph making the connection to the Malliavin calculus (Nourdin and Peccati [12]), and a treatment of all target laws in the Pearson class (Eden and Viens [5]). Let now  $X \sim \text{Poi}(\lambda)$ , the Poisson distribution, with probability mass function (pmf)

$$(2.5) \quad \mathbb{P}(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k \in \mathbb{Z}_+,$$

for some  $\lambda > 0$ . The Stein operator for  $X$  is given by (see Barbour *et al.* [3])

$$(2.6) \quad \mathcal{A}g(k) = \lambda g(k+1) - kg(k), \quad k \in \mathbb{Z}_+.$$

Also, the bounds for the solution of Stein equation (2.3) are given by

$$(2.7) \quad \|g\| \leq \frac{1}{\max(1, \sqrt{\lambda})} \quad \text{and} \quad \|\Delta g\| \leq \frac{2\|f\|}{\max(1, \lambda)},$$

where  $\Delta g(k) = g(k+1) - g(k)$  and  $\|\Delta g\| = \sup_k |g(k+1) - g(k)|$ . As noted in item 3 above, these bounds are essentially uniform for all functions  $f$ : the dependence on  $f$  is only via  $\|f\|$ . By the definition of the space of test functions  $\mathcal{H}$ , one may replace  $2\|f\|$  by 1 in (2.7) (see Upadhye *et al.* [16]), removing the dependence on  $f$  entirely. For more details, we refer the reader to Barbour *et al.* [3] and Upadhye *et al.* [16].

### 3. POISSON APPROXIMATION TO PSDs

We first consider the case of independent but non-identical PSDs. To implement our machinery, we choose to work with a specific parametric class of PSDs, where the elements which vary in the data are the parameters, not the function  $h$  defining the class which we fix. We allow for inhomogeneity “in time” and within each sequence of data points, by using a double (triangular) array of rvs. Specifically, let  $X_{i,n}$ ,  $1 \leq i \leq n$ ,  $n \geq 1$ , be a double array of independent rvs having PSDs with pmf

$$(3.1) \quad \mathbb{P}(X_{i,n} = k) = p_{i,n}(k) = \frac{a_k \theta_{i,n}^k}{h(\theta_{i,n})}, \quad k \in \mathbb{Z}_+,$$

where  $a_k \geq 0$ ,  $\theta_{i,n} > 0$ , and  $h(\theta_{i,n}) = \sum_{k=0}^{\infty} a_k \theta_{i,n}^k$ ,  $1 \leq i \leq n$ . We call, for simplicity,  $X_{i,n}$  power series rvs and the distribution in (3.1) the PSDs associated with the function  $h$ . We assume  $h$  is differentiable. Note that

$$(3.2) \quad \mathbb{E}X_{i,n} = \sum_{k=1}^{\infty} k p_{i,n}(k) = \frac{1}{h(\theta_{i,n})} \sum_{k=1}^{\infty} k a_k \theta_{i,n}^k = \frac{\theta_{i,n} h'(\theta_{i,n})}{h(\theta_{i,n})}.$$

Since  $h'(\theta_{i,n}) = \sum_{k=1}^{\infty} k a_k \theta_{i,n}^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} \theta_{i,n}^k$ , we have

$$(3.3) \quad \sum_{k=0}^{\infty} \frac{(k+1) a_{k+1} \theta_{i,n}^k}{h'(\theta_{i,n})} = 1.$$

Let  $X_{i,n}^*$ ,  $1 \leq i \leq n$ ,  $n \geq 1$ , be a sequence of independent power series rvs corresponding to  $h'$  so that pmf of  $X_{i,n}^*$  is given by

$$(3.4) \quad p_{i,n}^*(k) = \mathbb{P}(X_{i,n}^* = k) = \frac{(k+1)a_{k+1}\theta_{i,n}^k}{h'(\theta_{i,n})}, \quad k \in \mathbb{Z}_+.$$

Since  $h = h'$  in the case of the Poisson law, heuristically, one can ask oneself whether comparing the laws of the double arrays based on another  $h$  and its  $h'$  might tell us how close those laws are to Poisson ones. It turns out that this is a good strategy for partial sums, as our main theorem below shows, which explains why we introduce  $X^*$  above.

Specifically, for  $n \geq 1$ , let  $S_n = \sum_{i=1}^n X_{i,n}$  denote the sequence of the partial sums of the independent power series rvs. As announced in the introduction, our main interest in this article is to study Poisson approximation to  $S_n$ , and obtain error bounds. Our results unifies several known results obtained for specific distributions, such as binomial and geometric, belonging to PSDs.

Let  $N_\lambda$  henceforth denote the Poisson rv with mean  $\lambda > 0$ . As mentioned in the introduction, the next theorem is a device which shows that one only needs to keep track of two elements to approximate the law of  $S_n$  by that of  $N_\lambda$  in total variation. The first term in the approximation bound (3.14) below asks how close  $S_n$ 's mean is to  $\lambda$ . This is a minimal expression one could expect from moment matching. The second term is also a mean-type statistic, but can be thought of as a mixture mean expression, similar to what one encounters in Wald's identity. Here each summand in the mean of  $S_n$  is multiplied by a mean-type expression for the bounded variation measure of the difference between the law of the original  $X_{i,n}$  and of the derived  $X_{i,n}^*$  introduced above. As alluded to above, by depending explicitly on a mean-type statistic of  $p_{i,n} - p_{i,n}^*$ , which is small when  $h$  and  $h'$  are close to each other, this second term shows exactly how the distributional closeness of the summands to Poisson summands influences each term in the partial sum.

**THEOREM 3.1.** *Let  $X_{i,n}$ ,  $1 \leq i \leq n$ ,  $n \geq 1$ , be a double array of independent power series rvs defined in (3.1), and  $S_n = \sum_{i=1}^n X_{i,n}$ . Then*

$$(3.5) \quad d_{TV}(S_n, N_\lambda) \leq \frac{|\lambda - \mathbb{E}S_n|}{\max(1, \sqrt{\lambda})} + \frac{1}{\max(1, \lambda)} \sum_{i=1}^n \mathbb{E}X_{i,n} \sum_{k=1}^{\infty} k |p_{i,n}(k) - p_{i,n}^*(k)|,$$

where  $p_{i,n}(\cdot)$  and  $p_{i,n}^*(\cdot)$  are the PSDs given in (3.1) and (3.4), respectively.

**Proof.** Replacing  $k$  by  $S_n$  in (2.6) and taking expectation, we have

$$\mathbb{E}[\mathcal{A}g(S_n)] = \lambda \mathbb{E}[g(S_n + 1)] - \mathbb{E}[S_n g(S_n)].$$

Adding and subtracting  $\mathbb{E}S_n \mathbb{E}[g(S_n + 1)]$ , we get

$$(3.6) \quad \begin{aligned} \mathbb{E}[\mathcal{A}g(S_n)] &= (\lambda - \mathbb{E}S_n) \mathbb{E}[g(S_n + 1)] + \mathbb{E}S_n \mathbb{E}[g(S_n + 1)] - \mathbb{E}[S_n g(S_n)] \\ &= (\lambda - \mathbb{E}S_n) \mathbb{E}[g(S_n + 1)] + \sum_{i=1}^n \mathbb{E}X_{i,n} \mathbb{E}[g(S_n + 1)] \\ &\quad - \sum_{i=1}^n \mathbb{E}[X_{i,n} g(S_n)]. \end{aligned}$$

Let now  $W_{i,n} = S_n - X_{i,n}$  so that  $W_{i,n}$  and  $X_{i,n}$  are independent. Adding and subtracting  $\sum_{i=1}^n \mathbb{E}(X_{i,n} g(W_{i,n} + 1))$  in (3.6), we have

$$(3.7) \quad \begin{aligned} \mathbb{E}[\mathcal{A}g(S_n)] &= (\lambda - \mathbb{E}S_n) \mathbb{E}[g(S_n + 1)] + \sum_{i=1}^n \mathbb{E}X_{i,n} \mathbb{E}[g(S_n + 1) - g(W_{i,n} + 1)] \\ &\quad - \sum_{i=1}^n \mathbb{E}[X_{i,n} (g(S_n) - g(W_{i,n} + 1))]. \end{aligned}$$

First, consider the second term from (3.7)

$$\begin{aligned} \mathbb{E}[g(S_n + 1) - g(W_{i,n} + 1)] &= \mathbb{E}[g(W_{i,n} + X_{i,n} + 1) - g(W_{i,n} + 1)] \\ &= \sum_{k=0}^{\infty} \mathbb{E}[g(W_{i,n} + k + 1) - g(W_{i,n} + 1)] p_{i,n}(k), \end{aligned}$$

since  $W_{i,n}$  and  $X_{i,n}$  are independent. Note that

$$(3.8) \quad g(W_{i,n} + k + 1) - g(W_{i,n} + 1) = \sum_{j=1}^k \Delta g(W_{i,n} + j).$$

Therefore,

$$(3.9) \quad \mathbb{E}[g(S_n + 1) - g(W_{i,n} + 1)] = \sum_{k=1}^{\infty} \sum_{j=1}^k \mathbb{E}[\Delta g(W_{i,n} + j)] p_{i,n}(k).$$

Next, consider the third term of (3.7)

$$\begin{aligned}
\mathbb{E}[X_{i,n}(g(S_n) - g(W_{i,n} + 1))] &= \mathbb{E}[X_{i,n}(g(W_{i,n} + X_{i,n}) - g(W_{i,n} + 1))] \\
&= \sum_{k=0}^{\infty} k \mathbb{E}[g(W_{i,n} + k) - g(W_{i,n} + 1)] p_{i,n}(k) \\
&= \sum_{k=2}^{\infty} \sum_{j=1}^{k-1} k \mathbb{E}[\Delta g(W_{i,n} + j)] p_{i,n}(k) \quad (\text{using (3.8)}) \\
(3.10) \qquad \qquad \qquad &= \sum_{k=1}^{\infty} \sum_{j=1}^k (k+1) \mathbb{E}[\Delta g(W_{i,n} + j)] p_{i,n}(k+1).
\end{aligned}$$

Substituting (3.9) and (3.10) in (3.7), we get

$$\begin{aligned}
\mathbb{E}[\mathcal{A}g(S_n)] &= (\lambda - \mathbb{E}S_n) \mathbb{E}[g(S_n + 1)] + \sum_{i=1}^n \mathbb{E}X_{i,n} \sum_{k=1}^{\infty} \sum_{j=1}^k \mathbb{E}[\Delta g(W_{i,n} + j)] p_{i,n}(k) \\
&\quad - \sum_{i=1}^n \sum_{k=1}^{\infty} \sum_{j=1}^k (k+1) \mathbb{E}[\Delta g(W_{i,n} + j)] p_{i,n}(k+1) \\
&= (\lambda - \mathbb{E}S_n) \mathbb{E}[g(S_n + 1)] \\
&\quad + \sum_{i=1}^n \sum_{k=1}^{\infty} \mathbb{E}X_{i,n} \left[ p_{i,n}(k) - \frac{(k+1)p_{i,n}(k+1)}{\mathbb{E}X_{i,n}} \right] \sum_{j=1}^k \mathbb{E}[\Delta g(W_{i,n} + j)].
\end{aligned}$$

Therefore,

$$\begin{aligned}
(3.11) \quad |\mathbb{E}[\mathcal{A}g(S_n)]| &\leq |\lambda - \mathbb{E}S_n| \|g\| \\
&\quad + \|\Delta g\| \sum_{i=1}^n \sum_{k=1}^{\infty} k \mathbb{E}X_{i,n} \left| p_{i,n}(k) - \frac{(k+1)p_{i,n}(k+1)}{\mathbb{E}X_{i,n}} \right|.
\end{aligned}$$

Using (2.4), (2.7), and (3.11), we get

$$\begin{aligned}
(3.12) \quad d_{TV}(S_n, N_\lambda) &\leq \frac{|\lambda - \mathbb{E}S_n|}{\max(1, \sqrt{\lambda})} + \frac{1}{\max(1, \lambda)} \sum_{i=1}^n \mathbb{E}X_{i,n} \\
&\quad \times \sum_{k=1}^{\infty} k \left| p_{i,n}(k) - \frac{(k+1)p_{i,n}(k+1)}{\mathbb{E}X_{i,n}} \right|.
\end{aligned}$$

From (3.1) and (3.2), we have

$$(3.13) \quad \frac{(k+1)p_{i,n}(k+1)}{\mathbb{E}X_{i,n}} = \frac{(k+1)a_{k+1}\theta_{i,n}^k}{h'(\theta_{i,n})} = p_{i,n}^*(k).$$

Substituting (3.13) in (3.12), we get the required result. ■



The next corollary eliminates the first term in the bound simply by matching first moments. This is of practical interest because Theorem 3.1 is not an asymptotic result, holding globally for all  $n \geq 1$ , so that one may choose a value  $\lambda$  which depends on  $n$ , while the second term does not depend on  $\lambda$ , concerning itself only with individual summands' proximity to a Poisson law irrespective of their means.

**COROLLARY 3.1.** *Let  $X_{i,n}$ ,  $1 \leq i \leq n$ ,  $n \geq 1$ , be defined as in Theorem 3.1, and choose  $\lambda = \mathbb{E}S_n = \sum_{i=1}^n \mathbb{E}X_{i,n} = \sum_{i=1}^n \theta_{i,n} h'(\theta_{i,n})/h(\theta_{i,n})$  so that the first moment of  $N_\lambda$  and  $S_n$  match. Then*

$$(3.14) \quad d_{TV}(S_n, N_\lambda) \leq \frac{1}{\max(1, \lambda)} \sum_{i=1}^n \mathbb{E}X_{i,n} \sum_{k=1}^{\infty} k |p_{i,n}(k) - p_{i,n}^*(k)|.$$

Next we discuss some applications of Theorem 3.1 and Corollary 3.1. In the first example, in contrast to the previous corollary, one assumes that all data-points are Poisson distributed, leading to the disappearance of the second term.

**EXAMPLE 3.1.** *Let  $X_{i,n} \sim \text{Poi}(\lambda_{i,n})$ ,  $1 \leq i \leq n$  and  $n \geq 1$  so that  $a_k = 1/k!$ ,  $h(\lambda_{i,n}) = e^{\lambda_{i,n}}$ ,*

$$p_{i,n}(k) = \frac{e^{-\lambda_{i,n}} \lambda_{i,n}^k}{k!}, \quad k \in \mathbb{Z}_+;$$

$$p_{i,n}^*(k) = \frac{(k+1)a_{k+1}\lambda_{i,n}^k}{h'(\lambda_{i,n})} = \frac{e^{-\lambda_{i,n}} \lambda_{i,n}^k}{k!} = p_{i,n}(k), \quad k \in \mathbb{Z}_+.$$

*Then  $S_n = \sum_{i=1}^n X_{i,n} \sim \text{Poi}(\lambda_n)$ , where  $\lambda_n = \sum_{i=1}^n \lambda_{i,n}$ . From (3.5), we have*

$$(3.15) \quad d_{TV}(N_{\lambda_n}, N_\lambda) \leq \frac{|\lambda - \mathbb{E}S_n|}{\max(1, \sqrt{\lambda})} = \frac{|\lambda - \lambda_n|}{\max(1, \sqrt{\lambda})}.$$

*Also, if  $\lambda_n \rightarrow \lambda$ , as  $n \rightarrow \infty$ , then  $S_n \xrightarrow{\mathcal{L}} N_\lambda$ .*

**EXAMPLE 3.2.** *Let  $X_{i,n} \sim \text{Ber}(p_{i,n})$ ,  $1 \leq i \leq n$  so that  $p_{i,n}(k) = p_{i,n}^k (1 - p_{i,n})^{1-k}$ , for  $k = 0, 1$ . In that case,  $p_{i,n}^*(k) = 1$  for  $k = 0$  and zero otherwise. Also,  $a_k = 1$ , for  $k = 0, 1$ , and  $a_k = 0$  for  $k \geq 2$ ,  $h(\theta_{i,n}) = 1 + \theta_{i,n}$ , where  $\theta_{i,n} =$*

$p_{i,n}/(1 - p_{i,n})$ , and  $\mathbb{E}X_{i,n} = p_{i,n}$ . Then, from (3.5), we have

$$d_{TV}(S_n, N_\lambda) \leq \frac{|\lambda - \sum_{i=1}^n p_{i,n}|}{\max(1, \sqrt{\lambda})} + \frac{\sum_{i=1}^n p_{i,n}^2}{\max(1, \lambda)}.$$

Observe that  $d_{TV}(S_n, N_\lambda) \rightarrow 0$  if  $\sum_{i=1}^n p_{i,n} \rightarrow \lambda$ , and  $\sum_{i=1}^n p_{i,n}^2 \rightarrow 0$ , as  $n \rightarrow \infty$ , as proved in Theorem 3 of Wang [19]. This example thus provides a clear quantitative interpretation of Theorem 3 of Wang [19]. Further, if we allow  $\lambda$  to depend on  $n$  and use first-moment matching, then with  $\lambda = \sum_{i=1}^n p_{i,n}$ ,

$$(3.16) \quad d_{TV}(S_n, N_\lambda) \leq \frac{\sum_{i=1}^n p_{i,n}^2}{\max(1, \lambda)}.$$

The Poisson approximation to the sum of independent Bernoulli rvs has been studied by several authors and some bounds are given below:

- (i)  $d_{TV}(S_n, N_\lambda) \leq \sum_{i=1}^n p_{i,n}^2$  (Le Cam [11])
- (ii)  $d_{TV}(S_n, N_\lambda) \leq 1.05\lambda^{-1} \sum_{i=1}^n p_{i,n}^2$  (Kerstan [10])
- (iii)  $d_{TV}(S_n, N_\lambda) \leq (1 - e^{-\lambda})\lambda^{-1} \sum_{i=1}^n p_{i,n}^2$  (Barbour and Hall [2])

Note that we have used the bound for  $\|\Delta g\|$  given in (2.7) to obtain the bound in (3.16). We will get the Barbour and Hall [2] bound in (iii) if we use instead the bound  $\|\Delta g\| \leq (1 - e^{-\lambda})/\lambda$  (see (2.6) of Barbour and Hall [2]).

EXAMPLE 3.3. Let  $X_{i,n} \sim \text{Geo}(p_{i,n})$ ,  $1 \leq i \leq n$ . Then  $a_k = 1$ , for all  $k \in \mathbb{Z}_+$ ,  $h(\theta_{i,n}) = (1 - \theta_{i,n})^{-1}$ , where  $\theta_{i,n} = q_{i,n} = 1 - p_{i,n}$ . Note  $\mathbb{E}X_{i,n} = q_{i,n}/p_{i,n}$  and  $\mathbb{E}X_{i,n}^* = 2q_{i,n}/p_{i,n}$ . Also, when  $q_{i,n} \leq 1/2$ ,  $p_{i,n}^*(k) \geq p_{i,n}(k)$ ,  $k \in \mathbb{Z}_+$ . Therefore, from (3.14), we have for  $\lambda = \sum_{i=1}^n \frac{q_{i,n}}{p_{i,n}}$ ,

$$(3.17) \quad \begin{aligned} d_{TV}(S_n, N_\lambda) &\leq \frac{1}{\max(1, \lambda)} \sum_{i=1}^n \mathbb{E}X_{i,n} \sum_{k=1}^{\infty} k |p_{i,n}^*(k) - p_{i,n}(k)| \\ &= \frac{1}{\max(1, \lambda)} \sum_{i=1}^n \frac{q_{i,n}}{p_{i,n}} \left[ \frac{2q_{i,n}}{p_{i,n}} - \frac{q_{i,n}}{p_{i,n}} \right] = \frac{\sum_{i=1}^n \left( \frac{q_{i,n}}{p_{i,n}} \right)^2}{\max(1, \lambda)}, \end{aligned}$$

which is the bound obtained also by Kadu [9, p. 10]. Note that if  $\max_{1 \leq i \leq n} q_{i,n} \rightarrow 0$ , as  $n \rightarrow \infty$ , then  $S_n \xrightarrow{\mathcal{L}} N_\lambda$ .

Poisson approximation to the sum of independent (identical or non-identical) geometric rvs has been studied by several authors and the bounds are given below:

$$(i) \quad d_{TV}(S_n, N_\lambda) \leq (1 - e^{-\lambda}) \frac{q}{p}, \quad (\text{Barbour [1]})$$

$$(ii) \quad d_{TV}(S_n, N_\lambda) \leq \sum_{i=1}^n \frac{q_{i,n}^2}{p_{i,n}^2} \min \left\{ 1, \frac{1}{\sqrt{2\lambda e}} \right\}, \quad (\text{Vellaisamy and Upadhye [17]})$$

$$(iii) \quad \text{Using } |\mathbb{P}(S_n = k) - \mathbb{P}(N_\lambda = k)| \leq 2 \sum_{i=1}^n \left[ (1 - p_{i,n})^2 + \frac{1 - p_{i,n}}{p_{i,n}^2} \right],$$

(Hung and Giang [7]), for  $k \in \{0, 1, 2, \dots, n\}$ , we get a bound

$$(3.18) \quad \begin{aligned} d_{TV}(S_n, N_\lambda) &= \frac{1}{2} \sum_{k=0}^{\infty} |\mathbb{P}(S_n = k) - \mathbb{P}(N_\lambda = k)| \\ &\leq (n+1) \sum_{i=1}^n \left[ (1 - p_{i,n})^2 + \frac{1 - p_{i,n}}{p_{i,n}^2} \right] + \frac{1}{2} \sum_{k=n+1}^{\infty} |\mathbb{P}(S_n = k) - \mathbb{P}(N_\lambda = k)| \end{aligned}$$

Note that  $\frac{1}{\max(1, \lambda)} < n + 1$  and

$$\sum_{i=1}^n \left( \frac{q_{i,n}}{p_{i,n}} \right)^2 \leq \sum_{i=1}^n \left( q_{i,n}^2 + \frac{q_{i,n}}{p_{i,n}^2} \right) = \sum_{i=1}^n \left[ (1 - p_{i,n})^2 + \frac{1 - p_{i,n}}{p_{i,n}^2} \right].$$

Hence, the bound given in (3.17) is better than the bound given in (3.18).

$$(iv) \quad d_{TV}(S_n, N_\lambda) \leq \sum_{i=1}^n \min \left\{ \frac{\lambda^{-1}(1 - e^{-\lambda})}{p_{i,n}}, 1 \right\} q_{i,n}^2 p_{i,n}^{-1}, \quad (\text{Teerapabolarn and Wongkasem [15]})$$

Observe that the bound given in (3.17) is comparable to the bound given in (iv) which is an improvement over other bounds.

We next obtain a bound which may appear to be crude in comparison with Theorem 3.1, but it proves to be useful in situations where the PSDs parameters tend to 0. This is useful in applications, as shown in the last two examples in this section (e.g. geometric data with decaying likelihood of failure). This is a framework which was identified in Pérez-Abreu [13], where the corresponding theorem follows from our result, as explained below.

**THEOREM 3.2.** *Let  $X_{i,n}$ ,  $1 \leq i \leq n$ ,  $n \geq 1$ , be defined as in (3.1), where  $h$  is assumed to be twice differentiable. Let  $M_n = \sup_{\theta_{i,n}} [h'(\theta_{i,n})^2 + h''(\theta_{i,n})h(\theta_{i,n})]$ . If  $0 < M_n < \infty$ , then*

$$(3.19) \quad d_{TV}(S_n, N_\lambda) \leq \frac{|\lambda - \mathbb{E}S_n|}{\max(1, \sqrt{\lambda})} + \frac{M_n}{a_0^2 \max(1, \lambda)} \sum_{i=1}^n \theta_{i,n}^2.$$

*Proof.* Note that the right-hand side of (3.5) can be bounded by

$$(3.20) \quad d_{TV}(S_n, N_\lambda) \leq \frac{|\lambda - \mathbb{E}S_n|}{\max(1, \sqrt{\lambda})} + \frac{1}{\max(1, \lambda)} \sum_{i=1}^n \mathbb{E}X_{i,n} [\mathbb{E}X_{i,n} + \mathbb{E}X_{i,n}^*].$$

Observe that

$$(3.21) \quad \begin{aligned} \mathbb{E}X_{i,n}^* &= \sum_{k=1}^{\infty} k p_{i,n}^*(k) = \sum_{k=1}^{\infty} \frac{k(k+1)a_{k+1}\theta_{i,n}^k}{h'(\theta_{i,n})} \\ &= \frac{\theta_{i,n}}{h'(\theta_{i,n})} \sum_{k=1}^{\infty} k(k-1)a_k \theta_{i,n}^{k-2} = \frac{\theta_{i,n} h''(\theta_{i,n})}{h'(\theta_{i,n})}. \end{aligned}$$

Using (3.2) and (3.21) in (3.20), we get

$$(3.22) \quad \begin{aligned} d_{TV}(S_n, N_\lambda) &\leq \frac{|\lambda - \mathbb{E}S_n|}{\max(1, \sqrt{\lambda})} \\ &\quad + \frac{1}{\max(1, \lambda)} \sum_{i=1}^n \frac{\theta_{i,n} h'(\theta_{i,n})}{h(\theta_{i,n})} \left[ \frac{\theta_{i,n} h'(\theta_{i,n})}{h(\theta_{i,n})} + \frac{\theta_{i,n} h''(\theta_{i,n})}{h'(\theta_{i,n})} \right] \\ &= \frac{|\lambda - \mathbb{E}S_n|}{\max(1, \sqrt{\lambda})} \\ &\quad + \frac{1}{\max(1, \lambda)} \sum_{i=1}^n \left( \frac{\theta_{i,n}}{h(\theta_{i,n})} \right)^2 [h'(\theta_{i,n})^2 + h''(\theta_{i,n})h(\theta_{i,n})] \\ &\leq \frac{|\lambda - \mathbb{E}S_n|}{\max(1, \sqrt{\lambda})} + \frac{M_n}{\max(1, \lambda)} \sum_{i=1}^n \left( \frac{\theta_{i,n}}{h(\theta_{i,n})} \right)^2. \end{aligned}$$

Further, note that

$$(3.23) \quad h(\theta_{i,n}) = \sum_{k=0}^{\infty} a_k \theta_{i,n}^k \geq a_0 \implies \frac{1}{h(\theta_{i,n})^2} \leq \frac{1}{a_0^2}.$$

Using (3.23) in (3.22), the result follows. ■

First, we show that Theorem 3 of Pérez-Abreu [13] follows as a corollary.

COROLLARY 3.2. Let  $X_{i,n}$ ,  $1 \leq i \leq n$ ,  $n \geq 1$ , be defined as in (3.1) with  $a_0 > 0$  and  $S_n = \sum_{i=1}^n X_{i,n}$ . Also, assume

$$(3.24) \quad \theta_n^* = \max_{1 \leq i \leq n} \theta_{i,n} \rightarrow 0 \quad \text{and} \quad \sum_{i=1}^n \theta_{i,n} \rightarrow \lambda, \quad \text{as } n \rightarrow \infty,$$

for some  $\lambda > 0$ . Then  $S_n \xrightarrow{\mathcal{L}} N_{\lambda_0}$ , where  $\lambda_0 = \lambda a_1 / a_0$ .

PROOF. It suffices to show  $d_{TV}(S_n, N_{\lambda_0}) \rightarrow 0$ , as  $n \rightarrow \infty$ . First, we show that  $0 < M_n < \infty$ . Since  $h$  is an increasing function and  $\theta_n^* \rightarrow 0$ , as  $n \rightarrow \infty$ , we have  $a_0 \leq h(\theta_{i,n}) \leq h(\theta_n^*) \rightarrow a_0$ , showing that  $h(\theta_{i,n}) \rightarrow a_0$ , as  $\theta_n^* \rightarrow 0$ . Using the similar argument, since  $h'$  and  $h''$  are also increasing functions, we have  $h'(\theta_{i,n}) \rightarrow a_1$  and  $h''(\theta_{i,n}) \rightarrow 2a_2$ , as  $\theta_n^* \rightarrow 0$ . Also,

$$(3.25) \quad \begin{aligned} M_n &= \sup_{\theta_{i,n}} [h'(\theta_{i,n})^2 + h''(\theta_{i,n})h(\theta_{i,n})] \\ &\leq \sup_{\theta_n^*} [h'(\theta_n^*)^2 + h''(\theta_n^*)h(\theta_n^*)] \rightarrow (a_1^2 + 2a_0a_2), \end{aligned}$$

as  $\theta_n^* \rightarrow 0$ . Hence,  $0 < M_n < \infty$ .

Applying Theorem 3.2, we have from (3.19),

$$(3.26) \quad d_{TV}(S_n, N_{\lambda_0}) \leq \frac{|\lambda_0 - \mathbb{E}S_n|}{\max(1, \sqrt{\lambda_0})} + \frac{M_n \theta_n^*}{a_0^2 \max(1, \lambda_0)} \sum_{i=1}^n \theta_{i,n}.$$

Note also that

$$(3.27) \quad \mathbb{E}S_n = \sum_{i=1}^n \frac{\theta_{i,n} h'(\theta_{i,n})}{h(\theta_{i,n})} \rightarrow \frac{\lambda a_1}{a_0} = \lambda_0,$$

under the conditions (3.24). The result now follows from (3.26), (3.27) and the assumptions. ■

EXAMPLE 3.4. Let  $X_{i,n} \sim \text{Geo}(p_{i,n})$ ,  $1 \leq i \leq n$ . Then  $a_k = 1$ , for all  $k \in \mathbb{Z}_+$ ,  $h(\theta_{i,n}) = (1 - \theta_{i,n})^{-1}$ , where  $\theta_{i,n} = q_{i,n} = 1 - p_{i,n}$ . If  $\max_{1 \leq i \leq n} q_{i,n} \rightarrow 0$  and  $\sum_{i=1}^n q_{i,n} \rightarrow \lambda$ , as  $n \rightarrow \infty$ , then by Corollary 3.2,  $S_n \xrightarrow{\mathcal{L}} N_\lambda$ , since  $a_0 = a_1 = 1$ .

Next, we apply our results to the convolution of logarithmic series distributions which have many real-life applications. For example, Fisher *et al.* [6] used the logarithmic series distribution to investigate the distribution of butterflies in the

Malayan Peninsula. It is also used in areas such as the sampling of quadrants for plant species, the distribution of animal species, population and community ecology, population growth and some economic applications including inventory models, to mention a few.

The next example concerns the Poisson approximation to the sum of logarithmic series rvs, which has not been considered in the literature.

**EXAMPLE 3.5.** *Let  $Y_{i,n}$ ,  $1 \leq i \leq n$ ,  $n \geq 1$ , follow logarithmic series distribution with*

$$\mathbb{P}(Y_{i,n} = k) = -\frac{\theta_{i,n}^k}{k \ln(1 - \theta_{i,n})}, \quad 0 < \theta_{i,n} < 1, \quad k = 1, 2, 3, \dots$$

*Further, let  $X_{i,n} = Y_{i,n} - 1$ . Then*

$$(3.28) \quad \mathbb{P}(X_{i,n} = k) = p_{i,n}(k) = -\frac{\theta_{i,n}^{k+1}}{(k+1) \ln(1 - \theta_{i,n})}, \quad k \in \mathbb{Z}_+.$$

*Our interest is to obtain the error bound for the Poisson approximation to  $S_n = \sum_{i=1}^n X_{i,n}$ . Here,  $a_k = 1/(k+1)$ ,  $h(\theta_{i,n}) = -\ln(1 - \theta_{i,n})/\theta_{i,n}$  and therefore,*

$$(3.29) \quad p_{i,n}^*(k) = \frac{(k+1)a_{k+1}\theta_{i,n}^k}{h'(\theta_{i,n})} = \frac{(k+1)\theta_{i,n}^{k+1}}{(k+2) \left( \frac{1}{1-\theta_{i,n}} + \frac{\ln(1-\theta_{i,n})}{\theta_{i,n}} \right)}, \quad k \in \mathbb{Z}_+.$$

*From (3.14), we have*

$$(3.30) \quad d_{TV}(S_n, N_\lambda) \leq \frac{1}{\max(1, \lambda)} \sum_{i=1}^n \mathbb{E}X_{i,n} \sum_{k=1}^{\infty} k |p_{i,n}(k) - p_{i,n}^*(k)|,$$

*where  $\lambda = \sum_{i=1}^n \mathbb{E}X_{i,n}$ ,  $\mathbb{E}(X_{i,n}) = \mathbb{E}(Y_{i,n}) - 1 = -1 - \theta_{i,n}/[(1 - \theta_{i,n}) \ln(1 - \theta_{i,n})]$ , and  $p_{i,n}(k)$  and  $p_{i,n}^*(k)$  defined as in (3.28) and (3.29), respectively. Also, if  $\max_{1 \leq i \leq n} \theta_{i,n} \rightarrow 0$  and  $\sum_{j=1}^n \theta_{j,n} \rightarrow \lambda$ , as  $n \rightarrow \infty$ , then, by Corollary 3.2,  $S_n \xrightarrow{\mathcal{L}} N_{\lambda/2}$ , since  $a_0 = 1$  and  $a_1 = 1/2$ .*

**3.1. The Identical Distributions Case.** We end this section by applying our main results to the case of identically distributed data. This is a situation where the

PSDs parameter at level  $n$  must tend to zero as  $n \rightarrow \infty$ , to comply with obtaining a finite mean for the Poisson limit. We show how our theorems give fully quantitative versions of prior results in Pérez-Abreu [13], and new quantitative results for the Poisson approximations to the binomial and negative binomial distributions (see proofs of Corollaries 3.4 and 3.5.)

Let  $X_i$ ,  $1 \leq i \leq n$  ( $n \geq 1$ ), be a sequence of independent and identical rvs having the PSDs with pmf

$$(3.31) \quad p_n(k) = \mathbb{P}(X_i = k) = \frac{a_k \theta_n^k}{h(\theta_n)}, \quad k \in \mathbb{Z}_+,$$

where  $a_k \geq 0$ ,  $\theta_n > 0$ , and  $h(\theta_n) = \sum_{k=0}^{\infty} a_k \theta_n^k$ . Also, let  $X_i^*$ ,  $1 \leq i \leq n$ , be a sequence of independent rvs having PSDs corresponding to  $h'(\theta_n)$ , so that the pmf of  $X_i^*$  is given by

$$(3.32) \quad p_n^*(k) = \mathbb{P}(X_i^* = k) = \frac{(k+1)a_{k+1}\theta_n^k}{h'(\theta_n)}, \quad k \in \mathbb{Z}_+.$$

**THEOREM 3.3.** *Let  $X_i$  and  $X_i^*$ ,  $1 \leq i \leq n$  be a sequence of rvs with pmf's  $p_n(\cdot)$  and  $p_n^*(\cdot)$ , defined in (3.31) and (3.32), respectively. Let  $S_n = \sum_{i=1}^n X_i$ , and  $\mu_n = \mathbb{E}S_n = n\theta_n h'(\theta_n)/h(\theta_n)$ . Then, from Theorem 3.1, we have*

$$(3.33) \quad d_{TV}(S_n, N_\lambda) \leq \frac{|\lambda - \mu_n|}{\max(1, \sqrt{\lambda})} + \frac{\mu_n}{\max(1, \lambda)} \sum_{k=1}^{\infty} k |p_n(k) - p_n^*(k)|,$$

Also, from Theorem 3.2, a crude bound is

$$(3.34) \quad d_{TV}(S_n, N_\lambda) \leq \frac{|\lambda - \mu_n|}{\max(1, \sqrt{\lambda})} + \frac{n\theta_n^2 [h'(\theta_n)^2 + h''(\theta_n)h(\theta_n)]}{a_0^2 \max(1, \lambda)}.$$

The following result is Theorem 2 of Pérez-Abreu [13].

**COROLLARY 3.3.** *Let the conditions of Theorem 3.3 hold and  $n\theta_n \rightarrow \lambda$ , as  $n \rightarrow \infty$ , for some  $\lambda > 0$ . Then  $S_n \xrightarrow{\mathcal{L}} N_{\lambda_0}$ , where  $\lambda_0 = \lambda a_1/a_0$  with  $a_0 > 0$ .*

Next, we discuss two results, namely, binomial and negative binomial convergence to Poisson as application of Theorem 3.3.

**COROLLARY 3.4 (Poisson Approximation to Binomial Distribution).** *Let  $X_i \sim \text{Ber}(p_n)$ , for  $1 \leq i \leq n$ , and  $S_n = \sum_{i=1}^n X_i$ . If  $p_n \rightarrow 0$  and  $np_n \rightarrow \lambda$ , as  $n \rightarrow \infty$ , then  $S_n \xrightarrow{\mathcal{L}} N_\lambda$ .*

**Proof.** When  $X_i \sim \text{Ber}(p_n)$ ,  $1 \leq i \leq n$ , we have  $a_k = 1$ , for  $k = 0, 1$ , and  $a_k = 0$ , for  $k \geq 2$ . Also,  $h(\theta_n) = 1 + \theta_n$ , where  $\theta_n = p_n/(1 - p_n)$  and  $S_n \sim \text{Bi}(n, p_n)$ . Note that  $h'(\theta_n) = 1$  and  $h''(\theta_n) = 0$ . Hence, from (3.34),

$$(3.35) \quad d_{TV}(S_n, N_\lambda) \leq \frac{|\lambda - np_n|}{\max(1, \sqrt{\lambda})} + \frac{np_n^2}{\max(1, \lambda)(1 - p_n)^2}.$$

The bound given above goes to zero if  $p_n \rightarrow 0$  and  $np_n \rightarrow \lambda$ , as  $n \rightarrow \infty$ . This proves the result. ■

**COROLLARY 3.5 (Poisson Approximation to Negative Binomial Distribution).** *Let  $X_i \sim \text{Geo}(p_n)$ , for  $1 \leq i \leq n$ , and  $S_n = \sum_{i=1}^n X_i$ . If  $p_n \rightarrow 1$  and  $n(1 - p_n) \rightarrow \lambda$ , as  $n \rightarrow \infty$ , then  $S_n \xrightarrow{\mathcal{L}} N_\lambda$ .*

**Proof.** Since  $X_i \sim \text{Geo}(p_n)$ ,  $1 \leq i \leq n$ , we have  $a_k = 1$ , for all  $k \in \mathbb{Z}_+$ . Also,  $h(\theta_n) = (1 - \theta_n)^{-1}$ , where  $\theta_n = 1 - p_n$  and  $S_n \sim \text{NB}(n, p_n)$ . Note that  $h'(\theta_n) = (1 - \theta_n)^{-2} = p_n^{-2}$  and  $h''(\theta_n) = 2(1 - \theta_n)^{-3} = 2p_n^{-3}$ . Hence, from (3.34), we have

$$d_{TV}(S_n, N_\lambda) \leq \frac{|\lambda - \frac{n(1-p_n)}{p_n}|}{\max(1, \sqrt{\lambda})} + \frac{3n(1-p_n)^2}{\max(1, \lambda)p_n^4} \rightarrow 0,$$

if  $p_n \rightarrow 1$  and  $n(1 - p_n) \rightarrow \lambda$ , as  $n \rightarrow \infty$ . This proves the result. ■

**EXAMPLE 3.6.** Consider the similar argument discussed in Example 3.5 for logarithmic series distribution under identical setup. Then, we have  $a_k = 1/(k + 1)$  and  $h(\theta_n) = -\ln(1 - \theta_n)/\theta_n$ . Also, let  $n\theta_n \rightarrow \lambda$ , as  $n \rightarrow \infty$ . Then by Corollary 3.3,  $S_n \rightarrow N_{\lambda/2}$ , since  $a_0 = 1$  and  $a_1 = 1/2$ .

#### 4. COMPARISON OF POISSON AND NEGATIVE BINOMIAL BOUNDS

In this section, we compare the bounds for negative binomial approximation derived by Vellaisamy *et al.* [18] for the PSDs with those for Poisson approximation obtained in this paper through some relevant examples. Let  $M_{r,p}$  denote the



negative binomial rv with parameter  $r > 0$  and  $p = 1 - q \in (0, 1)$  with

$$\mathbb{P}(M_{r,p} = k) = \binom{r+k-1}{k} p^r q^k, \quad k \in \mathbb{Z}_+.$$

Also, let  $S_n = \sum_{i=1}^n X_{i,n}$ , where  $X_i$ 's are defined in (3.1) such that

$$(4.1) \quad \mathbb{E}S_n = \sum_{i=1}^n \mathbb{E}X_{i,n} = \sum_{i=1}^n \frac{\theta_{i,n} h'(\theta_{i,n})}{h(\theta_{i,n})} = \frac{rq}{p}.$$

The condition (4.1) implies  $\mathbb{E}(S_n) = \mathbb{E}(M_{r,p})$ , the first moment matching. Then, from Theorem 3.1 of Vellaisamy *et al.* [18], the one-parameter approximation bound is

$$\begin{aligned} d_{TV}(S_n, M_{r,p}) &\leq \frac{1}{rq} \sum_{i=1}^n \sum_{k=1}^{\infty} k |p \mathbb{E}X_{i,n} p_{i,n}(k) + qk p_{i,n}(k) - (k+1)p_{i,n}(k+1)| \\ &= \frac{1}{rq} \sum_{i=1}^n \mathbb{E}X_{i,n} \sum_{k=1}^{\infty} k \left| p p_{i,n}(k) + q \frac{k p_{i,n}(k)}{\mathbb{E}X_{i,n}} - \frac{(k+1)p_{i,n}(k+1)}{\mathbb{E}X_{i,n}} \right|. \end{aligned}$$

Using (3.13), we have

$$(4.2) \quad d_{TV}(S_n, M_{r,p}) \leq \frac{1}{rq} \sum_{i=1}^n \mathbb{E}X_{i,n} \sum_{k=1}^{\infty} k |p p_{i,n}(k) + q p_{i,n}^*(k-1) - p_{i,n}^*(k)|.$$

Furthermore, let

$$\begin{aligned} \tau &:= 2 \max_{1 \leq i \leq n} d_{TV}(W_{i,n}, W_{i,n} + 1) \\ &= \max_{1 \leq i \leq n} \sum_{k=1}^{\infty} |\mathbb{P}(W_{i,n} = k) - \mathbb{P}(W_{i,n} = k-1)|, \end{aligned}$$

where  $W_{i,n} = (S_n - X_{i,n})$ . Choose now  $r$  and  $p$  such that

$$(4.3) \quad r = \frac{(\mathbb{E}S_n)^2}{\text{Var}(S_n) - \mathbb{E}S_n} \quad \text{and} \quad p = \frac{\mathbb{E}S_n}{\text{Var}(S_n)}.$$

Then, from Theorem 4.1 of Vellaisamy *et al.* [18], two-parameter approximation bound is

$$\begin{aligned} d_{TV}(S_n, M_{r,p}) &\leq \frac{\tau}{rq} \sum_{i=1}^n \sum_{k=1}^{\infty} k \left( \frac{k-1}{2} + \mathbb{E}X_{i,n} \right) \\ &\quad \times |p \mathbb{E}X_{i,n} p_{i,n}(k) + qk p_{i,n}(k) - (k+1)p_{i,n}(k+1)|. \end{aligned}$$

Using (3.13), we have

$$(4.4) \quad d_{TV}(S_n, M_{r,p}) \leq \frac{\tau}{rq} \sum_{i=1}^n \mathbb{E}X_{i,n} \sum_{k=1}^{\infty} k \left( \frac{k-1}{2} + \mathbb{E}X_{i,n} \right) \\ \times |p p_{i,n}(k) + q p_{i,n}^*(k-1) - p_{i,n}^*(k)|.$$

Also, from Remark 4.1 of Vellaisamy *et al.* [18],  $\tau \leq \sqrt{\frac{2}{\pi}} \left( \frac{1}{4} + \sum_{i=1}^n \tau_i - \tau^* \right)^{-\frac{1}{2}}$ , where  $\tau_i = \min\{\frac{1}{2}, 1 - d_{TV}(X_{i,n}, X_{i,n} + 1)\}$  and  $\tau^* = \max_{1 \leq i \leq n} \tau_i$ .

When  $X_{i,n} \sim \text{Ber}(p_{i,n})$ ,  $1 \leq i \leq n$ , and fix  $r = n$  then, using Example 3.2 and (4.2), we have

$$(4.5) \quad d_{TV}(S_n, M_{n,p}) \leq \frac{2 \sum_{i=1}^n p_{i,n}^2}{\sum_{i=1}^n p_{i,n}},$$

where  $p = 1 / (1 + (1/n) \sum_{i=1}^n p_{i,n})$ . Observe that the Poisson approximation (see (3.16)) in this case is better than the negative binomial approximation. Also, the bound given in (4.4) is not valid as the condition (4.3) is not a valid choice of parameters.

When  $X_{i,n} \sim \text{Geo}(p_{i,n})$ ,  $1 \leq i \leq n$ , and for  $q_{i,n} \leq 1/2$ , we have from (14) and (19) of Vellaisamy *et al.* [18] that

$$(4.6) \quad d_{TV}(S_n, M_{r,p}) \leq \frac{1}{rq} \sum_{i=1}^n |p - p_{i,n}| \frac{q_{i,n}}{p_{i,n}^2}$$

and

$$(4.7) \quad d_{TV}(S_n, M_{r,p}) \leq 3 \sqrt{\frac{2}{\pi}} \left( \sum_{i=1}^n q_{i,n} - \frac{1}{4} \right)^{-1/2} \left( \sum_{i=1}^n \frac{q_{i,n}}{p_{i,n}} \right)^{-1} \sum_{i=1}^n \left| \frac{1}{p_i} - \frac{1}{p} \right| \left( \frac{q_{i,n}}{p_{i,n}} \right)^2,$$

respectively. Now, we give a numerical comparison of the above bounds with the Poisson approximation bound. Let us choose the values of  $q_{i,n} = 1 - p_{i,n}$  for various values of  $i$  as follows:

TABLE 1. The values of  $q_{i,n}$  for numerical computations.

$i$	1-10	11-20	21-50	51-100	101-150	151-200	201-250	251-300	301-400	401-500
$q_{i,n}$	0.20	0.18	0.16	0.14	0.12	0.10	0.08	0.06	0.04	0.02

Also, from (3.17), (4.6) and (4.7), we have the following numerical computation of the bounds.

TABLE 2. Comparison of bounds.

$n$	Poisson Approximation (using (3.17))	Negative Binomial Approximation (one moment matching, using (4.6))	Negative Binomial Approximation (two moments matching, using (4.7))
10	0.25	0	0
20	0.2357460	0.0152439	0.0045271
50	0.2108950	0.0221529	0.0040439
100	0.1897860	0.0242177	0.0029142
150	0.1754270	0.0279765	0.0028037
200	0.1638720	0.0329378	0.0025511
250	0.1543910	0.0379188	0.0025933
300	0.1468760	0.0436179	0.0025801
400	0.1365930	0.0531102	0.0024826
500	0.1312850	0.0612465	0.0024836

Note that, for  $n \leq 10$ ,  $p_{i,n} = 0.8$  and so the bounds for negative binomial approximation are zero as sum of iid geometric is negative binomial. From the table, we see that negative binomial approximation is better than Poisson approximation. Also, the bound obtained by matching first two moments is better than the bound obtained by matching one moment, as expected.

EXAMPLE 4.1. Consider again Example 3.5 for logarithmic series distribution. Note that

$$p_{i,n}(k-1) - p_{i,n}(k) = -\frac{\theta_{i,n}^k}{k(k+1)\ln(1-\theta_{i,n})} [k(1-\theta_{i,n}) + 1] \geq 0.$$

Therefore,

$$\begin{aligned} d_{TV}(X_{i,n}, X_{i,n} + 1) &= \frac{1}{2} \sum_{k=0}^{\infty} |p_{i,n}(k-1) - p_{i,n}(k)| \\ &= \frac{1}{2} \left[ p_{i,n}(0) + \sum_{k=1}^{\infty} [p_{i,n}(k-1) - p_{i,n}(k)] \right] \\ &= \frac{1}{2} [p_{i,n}(0) + p_{i,n}(0)] = p_{i,n}(0) = -\frac{\theta_{i,n}}{\ln(1-\theta_{i,n})}. \end{aligned}$$

Hence, the upper bounds in (4.2) and (4.4) are valid for logarithmic series distribution with  $\tau \leq \sqrt{\frac{2}{\pi}} \left( \frac{1}{4} + \sum_{i=1}^n \tau_i - \tau^* \right)^{-1/2}$ ,  $\tau_i = \min\{\frac{1}{2}, 1 + (\theta_{i,n}/\ln(1 - \theta_{i,n}))\}$  and  $\tau^* = \max_{1 \leq i \leq n} \tau_i$ ,  $\mathbb{E}(X_{i,n}) = \mathbb{E}(Y_{i,n}) - 1 = -1 - \theta_{i,n}/[(1 - \theta_{i,n}) \ln(1 - \theta_{i,n})]$ , and  $p_{i,n}(k)$  and  $p_{i,n}^*(k)$  defined as in (3.28) and (3.29), respectively. Also, for (4.2) and (4.4),  $r$  and  $p$  can be evaluated using the conditions (4.1) and (4.3), respectively.

From (3.30) and (4.2) with  $r = n/5$  and  $p = 1/(1 + (5/n) \sum_{i=1}^n (-1 - \theta_{i,n}/[(1 - \theta_{i,n}) \ln(1 - \theta_{i,n})]))$ , and (4.4), we have the following numerical computation of the bounds with  $\theta_{i,n} = q_{i,n}$  given in Table 1. It is difficult to compute the bounds in a compact form. So, the bounds are computed by using Mathematica Software, version 11.3.

TABLE 3. Comparison of bounds.

$n$	Poisson Approximation (from (3.30))	Negative Binomial Approximation (one moment matching, from (4.2))	Negative Binomial Approximation (two moments matching, from (4.4))
10	0.2068330	0.3949420	0.0033845
20	0.1950790	0.3711290	0.0033441
50	0.1745720	0.3293270	0.0028029
100	0.1571360	0.2931890	0.0022248
150	0.1452490	0.2661830	0.0019849
200	0.1356570	0.2411430	0.0019339
250	0.1277630	0.2165510	0.0019276
300	0.1214880	0.1917620	0.0019011
400	0.1128780	0.1479230	0.0018827
500	0.1084220	0.1086410	0.0018696

From the above table, it is clear that the Poisson approximation is better than negative binomial approximation with one moment matching. However, the negative binomial approximation with two moments matching is better than the Poisson approximation with one moment matching.

## ACKNOWLEDGEMENT

The authors are grateful to the associate editor and the reviewers for many valuable suggestions and critical comments that improved the presentation of the paper. A part of this work was carried out when the second author was visiting the Department of Statistics and Probability, Michigan State University, during Summer-2019.

## REFERENCES

- [1] A. D. Barbour. Asymptotic expansions in the Poisson limit theorem. *Ann. Probab.*, 15(2):748–766, 1987.
- [2] A. D. Barbour and P. Hall. On the rate of Poisson convergence. *Math. Proc. Cambridge Philos. Soc.*, 95(3):473–480, 1984.
- [3] A. D. Barbour, L. Holst, and S. Janson. *Poisson Approximation*. Oxford University Press, 1992.
- [4] L. H. Y. Chen. Poisson approximation for dependent trials. *Ann. Probab.*, 3(3):534–545, 1975.
- [5] R. Eden and F. Viens. General upper and lower tail estimates using Malliavin calculus and Stein’s equations. In Robert C. Dalang, Marco Dozzi, and Francesco Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications VII*, pages 55–84, Basel, 2013. Springer Basel.
- [6] R. A. Fisher, A. S. Corbet, and C. B. Williams. The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.*, 12(1):42–58, 1943.
- [7] T. L. Hung and L. T. Giang. On the bounds in Poisson approximation for independent geometric distribute random variables. *Bull. Iranian Math. Soc.*, 42(5):1087–1096, 2016.
- [8] N. L. Johnson, A. W. Kemp, and S. Kotz. *Univariate Discrete Distributions*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2005.
- [9] P. E. Kadu. Approximation results of sums of independent random variables. *To appear in Revstat Stat. J.*, 2020.
- [10] J. Kerstan. Verallgemeinerung eines Satzes von Prochorow und Le Cam. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 2:173–179, 1964.
- [11] L. Le Cam. An approximation theorem for the Poisson binomial distribution. *Pacific J. Math.*,

- 10:1181–1197, 1960.
- [12] I. Nourdin and G. Peccati. *Normal Approximations with Malliavin Calculus: from Stein's Method to Universality*. Cambridge Tracts in Mathematics. Cambridge Univ. Press, 2012.
- [13] V. Pérez-Abreu. Poisson approximation to power series distributions. *Am. Statist.*, 45(1):42–45, 1991.
- [14] K. Teerapabolarn. Poisson approximation for a sum of negative binomial random variables. *Bull. Malays. Math. Sci. Soc.*, 40(2):931–939, 2017.
- [15] K. Teerapabolarn and P. Wongkasem. Poisson approximation for independent geometric random variables. *Int. Math. Forum*, 2(65-68):3211–3218, 2007.
- [16] N. S. Upadhye, V. Čekanavičius, and P. Vellaisamy. On Stein operators for discrete approximations. *Bernoulli*, 23(4A):2828–2859, 2017.
- [17] P. Vellaisamy and N. S. Upadhye. Compound negative binomial approximations for sums of random variables. *Probab. Math. Statist.*, 29(2):205–226, 2009.
- [18] P. Vellaisamy, N. S. Upadhye, and V. Čekanavičius. On negative binomial approximation. *Theory Probab. Appl.*, 57(1):97–109, 2013.
- [19] Y. H. Wang. On the number of successes in independent trials. *Statist. Sinica*, 3(2):295–312, 1993.

Department of Mathematics  
Indian Institute of Technology Bombay  
Powai, Mumbai-400076, India.  
*E-mail*: amit.kumar2703@gmail.com

Department of Mathematics  
Indian Institute of Technology Bombay  
Powai, Mumbai-400076, India.  
*E-mail*: pv@math.iitb.ac.in

Department of Statistics and Probability  
Michigan State University  
East Lansing, MI 48824, USA.  
*E-mail*: viens@msu.edu

*Received on ;*  
*revised version on xx.xx.xxxx*

---