On perturbations of Stein operator

A. N. Kumar and N. S. Upadhye

Department of Mathematics, Indian Institute of Technology Madras, Chennai, Tamil Nadu, India

ABSTRACT
In this article, we obtain a Stein operator for the sum of \( n \) independent random variables (rvs) which is shown as the perturbation of the negative binomial (NB) operator. Comparing the operator with NB operator, we derive the error bounds for total variation distance by matching parameters. Also, three-parameter approximation for such a sum is considered and is shown to improve the existing bounds in the literature. Finally, an application of our results to a function of waiting time for \((k_1, k_2)\)-events is given.

1. Introduction
Applications of negative binomial (NB) distribution appear in many areas such as network analysis, epidemics, telecommunications, and related fields. NB approximation is widely studied in complex setting such as sum of waiting time, rare events, and extremes. Also, NB approximation to the sum of indicator random variables (rvs) is given by Brown and Phillips (1999), approximation of NB and NB perturbation to the sum of the independent rvs is given by Vellaisamy et al. (2013), NB approximation to the sum of independent NB rvs is given by Vellaisamy and Upadhye (2003), and NB approximation to \( k \)-runs is given by Wang and Xia (2008).

In this article, we obtain a Stein operator for the sum of independent rvs concentrated on \( \mathbb{Z}_+ = \{0, 1, 2, \ldots\} \). Stein operator, so obtained, is the perturbation of the NB operator. So, we investigate error in approximation for NB to the sum of independent rvs on \( \mathbb{Z}_+ \), via Stein method, by matching the first and second moments. Also, error in approximation for convolution of NB and a geometric rv to the sum of independent rvs on \( \mathbb{Z}_+ \) is investigated by matching the first three moments. An application of these investigations is demonstrated for a function of waiting time for \((k_1, k_2)\)-distribution. The approximation results, proved in Sections 3 and 4, are either comparable to or an improvement over the existing results in the literature.

Stein method (1986) is studied widely in probability approximations. For details and applications, see Barbour et al. (1992a), Holmes (2004), Goldstein and Reinert (2005), Daly (2010, 2011), Chen et al. (2011), Ross (2011), Daly et al. (2012), and Norudin and Peccati (2012). For recent developments, see Barbour and Chen (2014), Ley and Swan (2013a, 2013b), Ley et al. (2014), Upadhye et al. (2014), and references therein. This method involves
identifying a suitable operator (known as a Stein operator) which can be obtained using one of the approaches (see Reinert, 2005) such as density approach (Stein, 1986; Stein et al., 2004), generator approach (Barbour, 1990; Götz, 1991), and orthogonal polynomial approach (Diaconis and Zabell, 1991). Recently, a probability-generating function (PGF) approach (Upadhye et al., 2014) and a method to obtain a canonical Stein operator (Ley et al., 2014) are developed. We focus on PGF approach for finding Stein operators.

The article is organized as follows. In Section 2, we define some necessary notations to formulate Stein method and our main results. Also, we explain some known results for NB distribution from the literature. In Section 3, we first obtain a Stein operator for the sum of independent rvs which can be seen as perturbation of the NB operator. So, we obtain a bound between NB and the sum of independent rvs by matching one and two parameters. Next, we derive a Stein operator for convolution of NB and geometric rvs which motivates us to use perturbation technique for obtaining a bound between convolution of NB with a geometric and sum of independent rvs by matching three parameters. Finally, in Section 4, we give an application of our results for the function of waiting time for \((k_1, k_2)\) distribution.

2. Notations and known results

Throughout this article, let \(Z \sim NB(\alpha, p)\) with

\[
P(Z = m) = \binom{\alpha + m - 1}{m} p^\alpha q^m, \quad m = 0, 1, \ldots,
\]

where \(\alpha > 0\) and \(q = 1 - p \in (0, 1)\) and \(Y = \sum_{i=1}^{n} X_i\), where \(X_i, i = 1, 2, \ldots, n\) are independent rvs with PGF

\[
M_Y(z) := \mathbb{E}(z^Y) = \sum_{m=0}^{\infty} P(Y = m) z^m = \sum_{m=0}^{\infty} p_m z^m.
\]

Also, let the PGF of \(X_i\) be \(M_{X_i}\) such that

\[
G_{X_i}(z) := \frac{M'_{X_i}(z)}{M_{X_i}(z)} = \sum_{m=0}^{\infty} a_{i,m+1} z^m,
\]

and assume the series in (2) converges absolutely (see also, Yakshyavichus, 1998 for similar expansions of \(G_X(z)\)). In particular, for specific distributions, the following holds.

- (O1) \(X_i \sim Ge(p_i) \Rightarrow a_{i,m+1} = q_i^{m+1}\).
- (O2) \(X_i \sim Bi(\tilde{n}, \tilde{p}_i) \Rightarrow a_{i,m+1} = \tilde{n}(-1)^m (\tilde{p}_i/\tilde{q}_i)^{m+1}\).
- (O3) \(X_i \sim Po(\lambda_i) \Rightarrow a_{i,m+1} = \lambda_i\) for \(m = 0\) and 0 otherwise.

Next, let \(\mu\) and \(\sigma^2\) denote the mean and variance of \(Y\), respectively. Then

\[
\mu := \sum_{i=1}^{n} G'_{X_i}(1) = \sum_{i=1}^{n} \sum_{m=0}^{\infty} a_{i,m+1}, \quad \sigma^2 := \sum_{i=1}^{n} \sum_{m=0}^{\infty} (G_{X_i}(1) + G'_{X_i}(1)) = \sum_{i=1}^{n} \sum_{m=0}^{\infty} (m+1) a_{i,m+1},
\]

\[
\mu_2 := \sum_{i=1}^{n} G'_{X_i}(1) = \sum_{i=1}^{n} \sum_{m=0}^{\infty} ma_{i,m+1} \quad \text{and} \quad \mu_3 := \sum_{i=1}^{n} G''_{X_i}(1) = \sum_{i=1}^{n} \sum_{m=0}^{\infty} m(m-1) a_{i,m+1}.
\]

(3)
where $\mu_2$ and $\mu_3$ denote the second and third factorial cumulant moments of $Y$ (see Vellaisamy et al., 2013, pp. 104–105). Let us define

$$
\eta_1 := \frac{3}{2} \mu_2 \mu_3 - 4 \mu_2^2, \quad \eta_2 := 27 \mu_2^2 \mu_3^2 - 16 \mu_2^3 - \frac{27}{2} \mu_3^3 + 9 \mu_2 \mu_3 \mu_3, \\
\eta_3 := \left( \eta_2 + \sqrt{4 \eta_1^3 + \eta_3^2} \right)^{1/3} \quad \text{and} \quad \eta := 2 \mu_2 + \frac{\eta_1}{\eta_3} - \eta_3,
$$

provided $\eta, \eta_3 \in \mathbb{R}_+$, the set of all positive real numbers.

Now, let $G$ be the set of all bounded function on $\mathbb{Z}_+$ and

$$
G_X = \left\{ g | g \in G \text{ such that } g(0) = 0 \text{ and } g(x) = 0, \text{ for } x \notin \text{supp}(X) \right\}
$$

be associated with Stein operator $A_X$, where $\text{Supp}(X)$ denotes the support of a rv $X$.

Next, Stein method can be formulated in three steps. First, identify a suitable operator (known as a Stein operator) for the rv $X$. Stein operator is defined on family of function $G_X$ such that

$$
E(A_X g) = 0, \quad \text{for } g \in G_X.
$$

In the second step, we find the solution (say $g_f$) of the difference equation (known as Stein equation)

$$
A_X g(m) = f(m) - E f(Z), \quad m \in \mathbb{Z}_+ \text{ and } f \in G.
$$

and obtain the bound for $g_f$ (or $\Delta g_f$, as required) in terms of $f$.

Finally, substituting a rv $Y$ for $m$ in (6) and taking expectations and supremum, we get the following

$$
d_{TV}(Y, X) := \sup_{f \in \mathcal{H}} |E f(Y) - E f(X)| = \sup_{f \in \mathcal{H}} |E[A_X g_f(Y)]|,
$$

where $\mathcal{H} = \{ \mathbb{I}_S | S \text{ measurable} \}$ and $\mathbb{I}_S$ is the indicator function of the set $S$. Equation (7) is also equivalent to

$$
d_{TV}(Y, X) = \frac{1}{2} \sum_{m=0}^{\infty} |P(Y = m) - P(X = m)|.
$$

As $Y$ is the sum of independent rvs on $\mathbb{Z}_+$, from Corollary 1.6 of Mattner and Roos (2007), we have

$$
d_{TV}(Y, Y + 1) \leq \sqrt{\frac{2}{\pi}} \left( \frac{1}{4} + \sum_{i=1}^{n} (1 - d_{TV}(X_i, X_i + 1)) \right)^{-1/2}.
$$

Next, it is known that the Stein operator for NB($\alpha, p$) is given by (Brown and Phillips, 1999)

$$
(A_{Zg}) (m) = q(\alpha + m)g(m + 1) - mg(m), \quad \text{for } m \in \mathbb{Z}_+ \text{ and } g \in G_Z.
$$

Also, the bound for the solution to (6) is given by

$$
\|\Delta g\| \leq 1/aq.
$$

where $\|\Delta g\| = \sup_{m \in \mathbb{Z}_+} |\Delta g(m)|$ and $\Delta g(m) = g(m + 1) - g(m)$ denotes first forward difference operator (see Brown and Phillips, 1999 and Vellaisamy et al., 2013 for details).
As NB distribution can be described using two parameters, namely $\alpha$ and $p$, we can study the NB approximation problem using the moment matching technique up to the first two moments. For an approximation with extra parameter, we can use the perturbation technique described by Barbour et al. (2007) and can be formulated for NB distribution as follows:

Let $\mathcal{A}_{V}$ be a Stein operator of $V = Z + W$, where $W$ be a rv with parameter $\hat{p} = 1 - \hat{q}$ and probability mass function (PMF)

$$P(W = x) = \hat{q}^x \hat{p}, \quad x = 0, 1, 2, \ldots$$

Let $U_{V} = A_{V} - A_{Z}$, such that

$$||U_{V}g|| \leq \delta_{1} ||\Delta g||,$$

where $\alpha q > \delta_{1}$ and $g \in \mathcal{G}_{V}$, defined in (5). Also, let the rv $Y$ satisfy

$$|\mathbb{E}(A_{V}g)(Y)| \leq \delta_{2} ||\Delta g||, \quad \text{for } \delta_{2} \geq 0$$

then

$$d_{TV}(Y, V) \leq \frac{\delta_{2}}{\alpha q - \delta_{1}}. \quad (11)$$

(See Theorem 2.4 of Barbour et al., 2007 and (8) of Vellaisamy et al., 2013 for more general perturbation results.)

3. Approximation results

In this section, we obtain $Z$-approximation and $V$-approximation bounds to $Y$ using the first two moments and the first three moments respectively.

3.1. One-parameter approximation

The choice of the parameters can be done using the following relation.

$$\frac{\alpha q}{p} = \mu \quad \Rightarrow \quad p = \frac{\alpha}{\alpha + \mu} \quad \text{or} \quad \alpha = \frac{\mu p}{q}. \quad (12)$$

Here, matching can be done in two ways:

(i) Let $\alpha$ be fixed (in particular, $\alpha = n$) and $p = \alpha / (\alpha + \mu)$.

(ii) Let $p$ be fixed of our choice and the choice of $\alpha = \mu p/q$.

Theorem 3.1. Let $X_{1}, X_{2}, \ldots, X_{n}$ are independent rvs with (2) and $Y = \sum_{i=1}^{n} X_{i}$, then

$$d_{TV}(Y, Z) \leq \frac{1}{\alpha q} \sum_{i=1}^{n} \sum_{l=1}^{\infty} |a_{i, l+1} - qa_{i, l}|,$$

where $Z \sim NB(\alpha, p)$.

Proof. Given $Y = \sum_{i=1}^{n} X_{i}$ such that $X_{i}, i = 1, 2, \ldots, n$ are independent rvs. Then the PGF of $Y$ is given by $M_{Y}(z) = \prod_{i=1}^{n} M_{X_{i}}(z)$. Differentiating with respect to $z$, we have

$$M_{Y}'(z) = M_{Y}(z) \sum_{i=1}^{n} G_{X_{i}}(z) = \sum_{i=1}^{n} M_{Y}(z) \left( \sum_{m=0}^{\infty} a_{i, m+1} z^{m} \right),$$
where $G_X(\cdot)$ is defined in (2). Using (1) and multiplying by $1 - qz$, we get

$$\sum_{m=0}^{\infty} (m + 1) p_{m+1} z^m = q \sum_{m=0}^{\infty} (m + 1) p_{m+1} z^{m+1}$$

$$= \sum_{i=1}^{n} \left[ \sum_{m=0}^{\infty} \left( \sum_{l=0}^{m} p_l a_{i,m-l+1} \right) z^m - q \sum_{m=0}^{\infty} \left( \sum_{l=0}^{m-1} p_l a_{i,m-l} \right) z^m \right].$$

where $q = 1 - p$ and $p$ are defined in (12). Now, comparing the coefficient of $z^m$, we obtain the recursive relation

$$q m p_m - (m + 1) p_{m+1} + \sum_{i=1}^{n} \left( \sum_{l=0}^{m} p_l a_{i,m-l+1} - q \sum_{l=0}^{m-1} p_l a_{i,m-l} \right) = 0.$$

Let $g \in G_Y$, defined in (5), then

$$\sum_{m=0}^{\infty} g(m+1) \left[ q m p_m - (m + 1) p_{m+1} + \sum_{i=1}^{n} \left( \sum_{l=0}^{m} p_l a_{i,m-l+1} - q \sum_{l=0}^{m-1} p_l a_{i,m-l} \right) \right] = 0,$$

or equivalently

$$\sum_{m=0}^{\infty} \left[ q m g(m+1) - m g(m) + \left( \sum_{i=1}^{n} a_{i,1} \right) g(m+1) \right.$$

$$+ \sum_{i=1}^{n} \sum_{l=1}^{\infty} g(l + m + 1) \left( a_{i,l+1} - q a_{i,l} \right) \left] p_m = 0. \right.$$

Hence, the Stein operator for $Y$ is given by

$$A_Y g(m) = q m g(m+1) - m g(m) + \left( \sum_{i=1}^{n} a_{i,1} \right) g(m+1)$$

$$+ \sum_{i=1}^{n} \sum_{l=1}^{\infty} g(l + m + 1) \left( a_{i,l+1} - q a_{i,l} \right).$$

Rewrite Stein operator, using auxiliary parameter $\alpha > 0$, as

$$A_Y g(m) = q (\alpha + m) g(m+1) - m g(m) + \left( \sum_{i=1}^{n} a_{i,1} - \alpha q \right) g(m+1)$$

$$+ \sum_{i=1}^{n} \sum_{l=1}^{\infty} g(l + m + 1) \left( a_{i,l+1} - q a_{i,l} \right). \quad (13)$$

This is a Stein operator for the sum of independent rvs, which is a perturbation of NB$(\alpha, p)$ in view of Barbour and Xia (1999) and Vellaisamy et al. (2013). Applying Newton’s expansion as given in Barbour and Čekanavičius (2002), we have

$$g(m + l + 1) = \sum_{j=1}^{l} \Delta g(m + j) + g(m + 1). \quad (14)$$
Putting (14) in (13) and using (12), we get
\[
A_Y g(m) = q(\alpha + m)g(m + 1) - mg(m) + \sum_{i=1}^{\infty} \sum_{l=1}^{i} \sum_{j=1}^{l} \Delta g(m + j) (a_{i,l+1} - qa_{i,l})
\]
\[
= A_Z g(m) + U_Y g(m),
\]
where \(A_Z\) is a Stein operator for NB \((\alpha, p)\) described as in (9), \(A_Y\) is a Stein operator for sum of \(n\) independent rvs by matching mean with negative binomial rv. Now, for \(g \in G_Z \cap G_Y\), taking the expectation of \(U_Y\) with respect to \(Y\) and using (10), we get required result. \[\square\]

**Corollary 3.1.** Given \(Y = \sum_{i=1}^{n} X_i\), let \(X_i\) be the different types of distribution, we have the following bounds:

(i) Let \(X_i\) follow \(Ge(p_i)\), \(i = 1, 2, \ldots, n\) with \(q_i = (1 - p_i) < 1/2\), then
\[
d_{TV}(Y, Z) \leq \frac{1}{\alpha q} \sum_{i=1}^{n} \left| p - p_i \right| q_i^2 / q_i^2,
\]
where \(\sigma_X^2\) is the variance of \(X_i\).

(ii) Let \(X_i\) follow \(Po(\lambda_i)\) for \(i \in S_1\) and \(Ge(p_i)\) for \(i \in S_2\), where \(S_1 \cup S_2 = \{1, 2, \ldots, n\}\), then
\[
d_{TV}(Y, Z) \leq \frac{1}{\alpha q} \left( \sum_{i \in S_1} \lambda_i + \sum_{i \in S_2} \left| p - p_i \right| q_i / p_i^2 \right).
\]

(iii) Let \(X_i\) follow \(Bi(\tilde{n}, \tilde{p}_i)\) for \(i \in S_1\) and \(Ge(p_i)\) for \(i \in S_2\), where \(S_1 \cup S_2 = \{1, 2, \ldots, n\}\) with \(q_i, \tilde{p}_i < 1/2\), then
\[
d_{TV}(Y, Z) \leq \frac{1}{\alpha q} \left( \tilde{n} \sum_{i \in S_1} \left( \tilde{p}_i / q_i + q \right) \frac{\tilde{p}_i q_i}{(1 - 2\tilde{p}_i)^2} + \sum_{i \in S_2} \left| p - p_i \right| q_i^2 / p_i^2 \right).
\]

(iv) Let \(X_i\) follow \(Po(\lambda_i)\) for \(i \in S_1\) and \(Bi(\tilde{n}, \tilde{p}_i)\) for \(i \in S_2\), where \(S_1 \cup S_2 = \{1, 2, \ldots, n\}\) with \(\tilde{p}_i < 1/2\), then
\[
d_{TV}(Y, Z) \leq \frac{1}{\alpha q} \left( \sum_{i \in S_1} \lambda_i + \tilde{n} \sum_{i \in S_2} \left( \tilde{p}_i / q_i + q \right) \frac{\tilde{p}_i q_i}{(1 - 2\tilde{p}_i)^2} \right).
\]

**Remarks 3.1.**

(i) If \(a_{i,l+1} - qa_{i,l} \geq 0\) in Theorem 3.1, then using the definition of \(G_X\), it is easy to see that
\[
d_{TV}(Y, Z) \leq \frac{\sigma^2}{\mu - q}.
\]

(ii) The bound given in (15) is the same as the one given in (14), p. 101, of Vellaisamy et al. (2013), which is of constant order. Note that the approach used in proof is more general and easier than the approach used in Vellaisamy et al. (2013). Also, it is an improvement over Theorem 2.2 of Vellaisamy and Upadhye (2003) and comparable to Theorem 1 of Roos (2003).

(iii) If we replace \(p_i = p, i = 1, 2, \ldots, n\) in (15), then the bound is exact, as expected.

(iv) (16), (17), and (18) give bounds for the sum of two different types of rvs and can be easily extended for more than two different types of rvs.
(v) Instead of multiplying $(1 - qz)$ in the proof of Theorem 3.1, we can multiply the appropriate function to get the perturbation of some other known distribution, and hence the technique used can be generalized.

(vi) The bound, in (18), is not a good bound, as $Y$ has a mean greater than the variance but in NB the variance is bigger than the mean, as expected.

### 3.2. Two-parameter approximation

Next, we derive the bound between $Y$ and $Z$ by matching the first two moments, mean and variance, as

\[
\frac{\alpha q}{p} = \mu \text{ and } \frac{\alpha q}{p^2} = \sigma^2 \Rightarrow p = \frac{\mu}{\sigma^2} \text{ and } \alpha = \frac{\mu^2}{\mu_2},
\]

where $\mu, \mu_2,$ and $\sigma^2$ are defined in (3).

**Theorem 3.2.** Let $X_1, X_2, \ldots, X_n$ independent rvs with (2) and $Y = \sum_{i=1}^{n} X_i$, then

\[
d_{TV}(Y, Z) \leq \frac{1}{\alpha q} \sqrt{\frac{2}{\pi}} \left( \frac{1}{4} + \sum_{i=1}^{n} (1 - d_{TV}(X_i, X_{i+1})) \right)^{-1/2} \sum_{i=1}^{n} \sum_{l=1}^{\infty} l(l - 1)|a_{i,l+1} - qa_{i,l}|,
\]

where $\sigma^2 > \mu$ and $Z \sim \text{NB}(\alpha, p)$.

**Proof.** Using Newton’s expansion,

\[
\Delta g(m + j) = \sum_{u=1}^{j-1} \Delta^2 g(m + u) + \Delta g(m + 1).
\]

Substituting (14) and (20) in (13), we get

\[
A_Y g(m) = q(\alpha + m)g(m + 1) - mg(m)
\]

\[
+ \left[ \left( \sum_{i=1}^{n} a_{i,1} - \alpha q \right) + \sum_{i=1}^{n} \sum_{l=1}^{\infty} (a_{i,l+1} - a_{i,l}) \right] g(m + 1)
\]

\[
+ \sum_{i=1}^{n} \sum_{l=1}^{\infty} l(a_{i,l+1} - qa_{i,l}) \Delta g(m + 1)
\]

\[
+ \sum_{i=1}^{n} \sum_{l=1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta^2 g(m + u)(a_{i,l+1} - qa_{i,l}),
\]

where $q = 1 - p$ and $\alpha, p$ are defined in (19). Using (19) in (21), we obtain

\[
A_Y g(m) = q(\alpha + m)g(m + 1) - mg(m)
\]

\[
+ \sum_{i=1}^{n} \sum_{l=1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta^2 g(m + u)(a_{i,l+1} - qa_{i,l}) = A_Z g(m) + U_Y g(m).
\]

This is a Stein operator of sum of $n$ independent rvs by matching mean and variance with NB rv. Now, taking expectation of $U_Y$ w.r.t $Y$, we have
\[
\mathbb{E}[\tilde{U}_Y g(Y)] = \sum_{m=0}^{\infty} \left( \sum_{i=1}^{n} \sum_{l=1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{i-1} \Delta^2 g(m+u) (a_{i,l+1} - qa_{i,l}) \right) P(Y = m)
\]
\[
= \sum_{m=0}^{\infty} \sum_{i=1}^{n} \sum_{l=1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{i-1} \Delta g(m+u) \left[ P(Y = m - 1) - P(Y = m) \right] (a_{i,l+1} - qa_{i,l}).
\]

Therefore, for \( g \in \mathcal{G}_Z \cap \mathcal{G}_Y \), we have
\[
|\mathbb{E}[\tilde{U}_Y g(Y)]| \leq d_{TV}(Y, Y + 1) \| \Delta g \| \sum_{i=1}^{n} \sum_{l=1}^{\infty} l(l - 1) |a_{i,l+1} - qa_{i,l}|.
\]

Then, the proof is as follows by using (8) and (10).

\begin{corollary}
Given \( Y = \sum_{i=1}^{n} X_i \), let us choose \( X_i \) different type of distribution, then we have the following bounds:

(i) Let \( X_i \) follow Ge\((p_i)\), \( i = 1, 2, \ldots, n \) with \( q_i = (1 - p_i) < 1/2 \), then
\[
d_{TV}(Y, Z) \leq \frac{2}{\mu} \sqrt{\frac{2}{\pi}} \left( \frac{1}{4} + \sum_{i=1}^{n} \left( 1 - e^{-\lambda_i} \frac{\lambda_i^{[\lambda_i]}}{[\lambda_i]} \right) \right) + \sum_{i=1}^{n} q_i \right)^{-1/2} \left( \sum_{i=1}^{n} \left| p - p_i \right| \frac{q_i^2}{p_i^2} \right).
\]

(ii) Let \( X_i \) follow Po\((\lambda_i)\) for \( i \in S_1 \) and Ge\((p_i)\) for \( i \in S_2 \), where \( S_1 \cup S_2 = \{1, 2, \ldots, n\} \) and \( q_i < 1/2 \), then
\[
d_{TV}(Y, Z) \leq \frac{2}{\alpha q} \sqrt{\frac{2}{\pi}} \left( \frac{1}{4} + \sum_{i=1}^{n} \left( 1 - \frac{p_i^2 q_i}{(1 - 2q_i)^3} \right) + \sum_{i=1}^{n} q_i \right)^{-1/2} \left( \sum_{i=1}^{n} \left| p - p_i \right| \frac{q_i^2}{p_i^2} \right).
\]

(iii) Let \( X_i \) follow Bi\((\tilde{n}_i, \tilde{p}_i)\) for \( i \in S_1 \) and Ge\((p_i)\) for \( i \in S_2 \), where \( S_1 \cup S_2 = \{1, 2, \ldots, n\} \) with \( q_i, \tilde{q}_i < 1/2 \), then
\[
d_{TV}(Y, Z) \leq \frac{1}{\alpha q} \sqrt{\frac{2}{\pi}} \left( \frac{\tilde{n}_i \sum_{i \in S_1} \left( \frac{\tilde{p}_i}{\tilde{q}_i} + q \right) + \tilde{p}_i^2 \tilde{q}_i}{(1 - 2\tilde{q}_i)^3} + 2 \sum_{i \in S_2} \left| p - p_i \right| \frac{q_i^2}{p_i^2} \right)^{1/2},
\]
\[
where \frac{\tilde{n}_i \sum_{i \in S_1} \tilde{p}_i^2}{\tilde{q}_i^2} \leq \sum_{i \in S_2} \frac{q_i^2}{p_i^2}.
\]
\end{corollary}

\begin{remarks}
(i) If \( p_i = p, i = 1, 2, \ldots, n \) in (22), then the bound is exact, as expected.
(ii) The bound in (22) is improvement by constant over Corollary 4.1 of Vellaisamy et al. (2013), which is of order \( O(n^{-1/2}) \).
(iii) (23) and (24) give the bound for two different types of rvs, where the variance is greater than the mean. Also, this can be extended for more than two rvs.
\end{remarks}
3.3. Three-parameter approximation

As mentioned in Section 2, NB distribution can be described using two parameters. Therefore, for three-parameter approximation, we use convolution of one-parameter distribution, namely geometric, with NB. Convolution of Poisson with NB is studied by Vellaisamy et al. (2013) and has improved the accuracy of approximation with respect to NB or Poisson approximation. Therefore, we choose the geometric distribution as it has a behavior similar to the NB distribution.

Next, we derive a Stein operator for the convolution of NB and geometric rv. Recall that $Z \sim \text{NB}(\alpha, p)$ and $W \sim \text{Ge}(\hat{p})$, then the PGF of $Z$ and $W$ is given by $M_Z(z) = p^\alpha/(1 - qz)^\alpha$ and $M_W(z) = \hat{p}/(1 - \hat{q}z)$ respectively. Also, $V = W + Z$, then the PGF of $V$ is $M_V(z) = p^{\alpha} \hat{p}/((1 - qz)^\alpha(1 - \hat{q}z))$. Differentiating with respect to $z$ and multiplying by $(1 - qz)$, we get

$$
\sum_{m=0}^{\infty} (m + 1) p'_{m+1} z^m - q \sum_{m=0}^{\infty} m p'_{m} z^m = \alpha q \sum_{m=0}^{\infty} p'_{m} z^m \\
+ \sum_{m=0}^{\infty} \left( \sum_{l=0}^{m} p'_{l} \hat{q}^{m-l+1} \right) z^m - q \sum_{m=0}^{\infty} \left( \sum_{l=0}^{m-1} p'_{l} \hat{q}^{m-l} \right) z^m.
$$

where $p'_m = P(V = m)$ be the PMF of $V$. Comparing the coefficient of $z^m$, we have

$$
\alpha q p'_m + q m p'_{m} - (m + 1) p'_{m+1} = q \sum_{l=0}^{m-1} p'_{l} \hat{q}^{m-l} - \sum_{l=0}^{m} p'_{l} \hat{q}^{m-l+1} = \sum_{l=0}^{m} (\hat{q} - q) p'_{l} \hat{q}^{m-l} - q p'_m,
$$

This can be written as

$$
q(\alpha + 1 + m) p'_m - (m + 1) p'_{m+1} + \sum_{l=0}^{m} (\hat{q} - q) p'_{l} \hat{q}^{m-l} = 0.
$$

For $g \in G_V$, defined in (5), we have

$$
\sum_{m=0}^{\infty} g(m+1) \left[ (\alpha + 1) q p'_m + q m p'_{m} - (m + 1) p'_{m+1} + \sum_{l=0}^{m} (\hat{q} - q) p'_{l} \hat{q}^{m-l} \right] = 0
$$

Hence,

$$
\mathbb{E}[A \hat{V} g(V)] = \sum_{m=0}^{\infty} \left[ q(\alpha + 1 + m) g(m+1) - mg(m) + (\hat{q} - q) \sum_{l=0}^{\infty} g(m + l + 1) \hat{q}^{l} \right] p'_m = 0,
$$

where $A \hat{V} g(m) = q(\alpha + 1 + m) g(m+1) - mg(m) + (\hat{q} - q) \sum_{l=0}^{\infty} g(m + l + 1) \hat{q}^{l}$ is a Stein operator for $V$, which is a perturbation of NB$(\alpha + 1, p)$. Using (14), the Stein operator can be written as

$$
A \hat{V} g(m) = q \left[ (\alpha + 1 + \frac{\hat{q} - q}{q \hat{p}}) + m \right] g(m+1) - mg(m) \\
+ \left( \frac{\hat{q} - q}{\hat{p}} \right) \sum_{j=1}^{\infty} \Delta g(m + j) \hat{q}^{j} = \hat{A} \hat{Z} g(m) + \hat{U} \hat{V} g(m),
$$

where $\hat{A} \hat{Z}$ is a Stein operator for NB$(r, p)$ with $r = (\alpha + 1 + \frac{\hat{q} - q}{q \hat{p}})$. Then

$$
\mathbb{E}[\hat{U} \hat{V} g(V)] \leq \| \Delta g \| \times | \hat{q} - q| \frac{\hat{q}}{p}.
$$
Next, we match the first three moments of \( V \) and \( Y \), we have
\[
\frac{\alpha q}{\hat{p}} + \frac{\hat{q}^2}{2\hat{p}} = \mu, \quad \frac{\alpha q^2}{\hat{p}^2} + \frac{\hat{q}^3}{3\hat{p}^2} = \mu_2, \quad \text{and } \frac{\alpha q^3}{\hat{p}^3} + \frac{\hat{q}^4}{4\hat{p}^3} = \frac{\mu_3}{2},
\]
(26)
where \( \mu, \mu_2, \) and \( \mu_3 \) are defined in (3). Therefore, the choice of parameters is
\[
\hat{p} = \frac{3\mu}{3\mu + \eta}, \quad \alpha = \left( \mu - \frac{\eta}{3\mu} \right)^2 / \left( \mu_2 - \frac{\eta^2}{9\mu^2} \right) \text{ and } p
\]
(27)
where \( \eta \) is defined in (4). Now, we obtain the bound for \( V \) approximation to \( Y \) by matching the first three moments.

**Theorem 3.3.** Let \( X_1, X_2, \ldots, X_n \) be independent rvs with (2) and \( \sigma^2 > \mu \), then
\[
d_{TV}(Y, V) \leq \frac{16}{\Psi \times \left( rq - |\hat{q} - q| \frac{\hat{q}}{\hat{p}^2} \right)} \left( \sum_{i=1}^{k} \sum_{l=1}^{\infty} \frac{l(l-1)(l-2)}{6} |a_{i,l+1} - qa_{i,l}| + \frac{|\hat{q} - q| \hat{q}^3}{\hat{p}^4} \right),
\]
where \( \Psi = \sum_{i=1}^{n} \xi_i, \xi_i = \min_{1 \leq i \leq n}\left( \frac{\xi}{2}, 1 - d_{TV}(X_i, X_i + 1) \right) \) and \( rq > |\hat{q} - q| \frac{\hat{q}}{\hat{p}} \).

**Proof.** Now, for \( g \in \mathcal{G}_r \), we introduce a parameter \( \hat{q} \) and modify the Stein operator of \( Y \) in (21) as follows:
\[
\hat{A}_Y g(m) = q(\alpha + 1 + m)g(m + 1) - mg(m) + (\hat{q} - q) \sum_{l=0}^{\infty} g(m + l + 1) \hat{q}^l
\]
\[
+ \left( \sum_{i=1}^{n} a_{i,1} - \alpha q - q \right) + \sum_{i=1}^{n} \sum_{l=1}^{\infty} \left( a_{i,l+1} - qa_{i,l} \right) g(m + 1)
\]
\[
+ \sum_{i=1}^{n} \sum_{l=1}^{\infty} l \left( a_{i,l+1} - qa_{i,l} \right) \Delta g(m + 1)
\]
\[
+ \sum_{i=1}^{n} \sum_{l=1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta^2 g(m + u) \left( a_{i,l+1} - qa_{i,l} \right) - (\hat{q} - q) \sum_{l=0}^{\infty} g(m + l + 1) \hat{q}^l
\]
\[
= \hat{A}_Y g(m) + \hat{U}_Y g(m),
\]
where \( q = 1 - p \) and \( \hat{q} = 1 - \hat{p} \). Also, \( \alpha, p, \) and \( \hat{p} \) are defined in (27). Again, from Newton’s expansion, we have
\[
\Delta^2 g(m + u) = \sum_{v=1}^{u-1} \Delta^3 g(m + v) + \Delta^2 g(m + 1).
\]
(28)
Substituting (14), (20), and (28) in \( \hat{U}_Y \) and then using (26), we get
\[
\hat{U}_Y g(m) = \sum_{i=1}^{n} \sum_{l=1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \sum_{v=1}^{u-1} \Delta^3 g(m + v) \left( a_{i,l+1} - qa_{i,l} \right) - \frac{(\hat{q} - q) \hat{q}^3}{\hat{p}^3} \sum_{v=1}^{\infty} \Delta^3 g(m + v) \hat{q}^v.
\]
(29)
Taking expectation w.r.t $Y$, we have

$$
\mathbb{E}(\hat{U}_1 g(Y)) = \sum_{m=0}^{\infty} \left( \sum_{i=1}^{n} \sum_{l=1}^{\infty} \sum_{j=1}^{l-1} \sum_{u=1}^{j-1} \sum_{v=1}^{u-1} \Delta^3 g(m + v) \left( a_{i,l+1} - qa_{i,l} \right) \\
- \frac{\hat{q} - q}{p^3} \sum_{v=1}^{\infty} \Delta^3 g(m + v) \hat{q}^v \right) \rho_m
$$

$$
= \sum_{m=0}^{\infty} \left( \sum_{i=1}^{n} \sum_{l=1}^{\infty} \sum_{j=1}^{l-1} \sum_{u=1}^{j-1} \sum_{v=1}^{u-1} \Delta g(m + v) \{ p_{m-2} - 2p_{m-1} + p_m \} \left( a_{i,l+1} - qa_{i,l} \right) \\
- \frac{\hat{q} - q}{p^3} \sum_{v=1}^{\infty} \Delta g(m + v) \{ p_{m-2} - 2p_{m-1} + p_m \} \hat{q}^v \right)
$$

Hence,

$$
\left| \mathbb{E}(\hat{U}_1 g(Y)) \right| \leq \| \Delta g \| \left( \sum_{i=1}^{n} \sum_{l=1}^{\infty} \frac{l(l-1)(l-2)}{6} \left| a_{i,l+1} - qa_{i,l} \right| + \frac{|\hat{q} - q|}{p^3} \right) \sum_{m=0}^{\infty} \| p_{m-2} - 2p_{m-1} + p_m \|.
$$

(30)

From (4.9) of Barbour and Čekanavičius (2002), we have

$$
\sum_{m=0}^{\infty} \| p_{m-2} - 2p_{m-1} + p_m \| \leq \frac{16}{\Psi},
$$

where $\Psi = \sum_{i=1}^{k} \xi_i$ and $\xi_i = \min_{1 \leq i \leq n} \left\{ \frac{1}{2}, \frac{d_{TV}(X_i, X_i + 1)}{2} \right\}$.

Hence, from (11) and (25), the proof follows.$\square$

**Corollary 3.3.** Let $Y = \sum_{i=1}^{n} X_i$ such that $X_i$ be the different types of distribution, then for $rq > |\hat{q} - q| \frac{\hat{q}}{p^2}$, we have the following bounds:

(i) Let $X_i$ follow Ge($p_i$), $i = 1, 2, \ldots, n$ with $q_i = (1 - p_i) < 1/2$, then

$$
d_{TV}(Y, V) \leq \frac{16 (\sum_{i=1}^{n} q_i)^{-1}}{p \times \left( rq - |\hat{q} - q| \frac{\hat{q}}{p^2} \right)} \left( \sum_{i=1}^{n} \left| \frac{1}{p} - \frac{1}{p_i} \right| \left( \frac{q_i}{p_i} \right)^3 + \frac{1}{p} - \frac{1}{p} \left| \frac{\hat{q}}{p} \right| \left( \frac{\hat{q}}{p} \right)^3 \right).
$$

(31)

(ii) Let $X_i$ follow Po($\lambda_i$) for $i \in S_1$ and Ge($p_i$) for $i \in S_2$, where $S_1 \cup S_2 = \{ 1, 2, \ldots, n \}$ and $q_i < 1/2$, then

$$
d_{TV}(Y, V) \leq \frac{16 \Psi^{-1}}{p \times \left( rq - |\hat{q} - q| \frac{\hat{q}}{p^2} \right)} \left( \sum_{i \in S_2} \left| \frac{1}{p} - \frac{1}{p_i} \right| \left( \frac{q_i}{p_i} \right)^3 + \frac{1}{p} - \frac{1}{p} \left| \frac{\hat{q}}{p} \right| \left( \frac{\hat{q}}{p} \right)^3 \right),
$$

(32)
(iii) Let $X_i$ follow $Bi(\tilde{n}, \tilde{p}_i)$ for $i \in S_1$ and $Ge(p_i)$ for $i \in S_2$, where $S_1 \cup S_2 = \{1, 2, \ldots, n\}$ with $q_i, \tilde{p}_i < 1/2$, then

$$d_{TV}(Y, V) \leq \frac{16}{\Psi} \sum_{i \in S_1} \left( \frac{\tilde{p}_i q_i + q}{1 - 2 \tilde{p}_i} \right)^2 + \sum_{i \in S_2} \left| \frac{1}{p_i} - \frac{1}{\tilde{p}_i} \right| \left( \frac{q_i}{p_i} \right)^3 + \frac{1}{\Psi} \left( \frac{\hat{q}}{p} \right)^3,$$

where $\tilde{n} \sum_{i \in S_1} \tilde{p}_i^2 < \sum_{i \in S_2} \frac{q_i^2}{\tilde{p}_i^2}$.

Remarks 3.3.

(i) If $p_i = p$, $i = 1, 2, \ldots, n$ and $\tilde{p} = p$ in (31), then the bound is exact, as expected.

(ii) The bound in Theorem 3.3 is of the order $O(n^{-1})$, which is improvement over one- and two-parameter approximation.

(iii) (31) and (32) give bounds for one type of rvs and two different type of rvs. Also, this can extend to more than two types of rvs.

(iv) We cannot obtain the bound for sums of binomial and Poisson rvs for two- and three-parameter approximation because the mean is greater than the variance. So, the choice of parameters is inadmissible.

4. An application

In this section, we demonstrate an application of our approximation results to obtain the bound between NB and a function of waiting time for binomial distribution of order $(k_1, k_2)$ (see Huang and Tsai, 1991).

Let $S$ denote success and $F$ failure, with success probability $\tilde{p}$, in a sequence of independent Bernoulli trials. If $k_1$ consecutive $F$s followed by $k_2$ consecutive $S$s, i.e.,

$$\ldots F \ldots F S \ldots S \ldots,$$

occurred then it is called $(k_1, k_2)$ event, where $(k_1, k_2)$ is a pair of non negative integers, including 0, excluding $(0,0)$. Also, let $\tilde{N}(n; k_1, k_2)$ be the number of occurrences of $(k_1, k_2)$ events in $n$ trials. The distribution of $\tilde{N}(n; k_1, k_2)$, denoted by $p_{x,n}$, is called the binomial distribution of order $(k_1, k_2)$. $p_{x,n}$ defined in Lemma 1 of Huang and Tsai (1991) as follows:

Lemma 4.1.

(i) $p_{x,n} = \begin{cases} 0 & \text{if } n < k_1 + k_2, \ x > 0; \\ 1 & \text{if } n < k_1 + k_2, \ x = 0; \\ \hat{q}^{k_1} \tilde{p}^{k_2} & \text{if } n = k_1 + k_2, \ x = 1; \\ 1 - \hat{q}^{k_1} \tilde{p}^{k_2} & \text{if } n = k_1 + k_2, \ x = 0. \end{cases}

(ii) $p_{0,n} = \begin{cases} 0 & \text{if } n < k_1 + k_2, \ x = 0; \\ 1 & \text{if } n < k_1 + k_2, \ x = 0; \\ \hat{q}^{k_1} \tilde{p}^{k_2} p_{0,n-k_1-k_2}. \end{cases}

(iii) $p_{x,n} = \sum_{j=0}^{\lfloor n / (k_1 + k_2) \rfloor} \hat{q}^{k_1} \tilde{p}^{k_2} p_{x-1,n} p_{0,n-k_1-k_2}.

(iv) $p_{x,n+1} = p_{x,n} + \hat{q}^{k_1} \tilde{p}^{k_2} \left[ p_{x-1,n-k_1-k_2+1} - p_{x,n-k_1-k_2+1} \right]$ for $n \geq k_1 + k_2, \ 1 \leq x \leq \left\lfloor \frac{n}{k_1 + k_2} \right\rfloor$,

where $\lfloor a \rfloor$ denote the greatest integer not exceeding $a$. 


Next, let \( \tilde{T}_n \) denote the waiting time for \( n \)th occurrence of \((k_1, k_2)\) event. Then
\[
\tilde{T}_n = T_1 + T_2 + \cdots + T_n,
\]
where \( T_j \) is \( k := k_1 + k_2 \) plus the number of trials between the \((j - 1)\)th and \( j \)th occurrence of \((k_1, k_2)\) event. \( T_j \)'s are independent and identically distributed (i.i.d.) with i.i.d. copy \( T \) having PMF
\[
P(T = n) = \begin{cases} 
0 & n < k; \\
\alpha(\bar{p}) & n = k; \\
\alpha(\bar{p})p_{0,n-k} & n > k,
\end{cases}
\]
where \( \alpha(\bar{p}) = \bar{q}^{k_1} \bar{p}^{k_2} \).

Define \( \hat{T}_j = T_j - k \), for \( j = 1, 2, \ldots, n \). Therefore, \( \hat{T}_j \) is the number of trials between \((j - 1)\)th and \( j \)th occurrence of \((k_1, k_2)\) event. Suppose \( \hat{T} \) be the i.i.d. copy of \( \hat{T}_j \). Then
\[
P(\hat{T} = n) = P(T = n + k) = \begin{cases} 
0 & n < 0; \\
\alpha(\bar{p}) & n = 0; \\
\alpha(\bar{p})p_{0,n} & n > 0.
\end{cases}
\]

Let \( M_{\hat{T}}(t) \) be the PGF of \( \hat{T} \). Then, it can be easily seen that

\[
M_{\hat{T}}(z) = \frac{\alpha(\bar{p})}{1 - z + \alpha(\bar{p})z^k}
\]
(see Huang and Tsai, 1991, pp. 128–129, for details). Define \( \hat{T} \) as, the number of failures before \( n \)th occurrence of \((k_1, k_2)\) event,
\[
\hat{T} = \hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_n.
\]
Then, the PGF of \( \hat{T} \) is
\[
M_{\hat{T}}(z) = \left( \frac{\alpha(\bar{p})}{1 - z + \alpha(\bar{p})z^k} \right)^n.
\]

Also, define
\[
b_{m,\bar{p}} := \sum_{l=0}^{[m/k]} (-1)^l \binom{m-l(k-1)}{l} a(\bar{p})^l.
\]

For more details of \((k_1, k_2)\) distribution, we refer the readers to Huang and Tsai (1991), Balakrishnan and Koutras (2002), and Dafnis et al. (2010).

### 4.1. One-parameter approximation

First, we derive the bound between \( \hat{T} \) and \( Z \) by matching the first moment as follows:

\[
\frac{\alpha q}{p} = n \left( \frac{1 - k\alpha(\bar{p})}{\alpha(\bar{p})} \right).
\]

Here, matching can be done in two ways:

1. Let \( p \) be fixed, of our choice, and \( \alpha = np(1 - k\alpha(\bar{p}))/qa(\bar{p}) \).
2. Let \( \alpha \) be fixed and \( p = \alpha(\bar{p})/(\alpha a(\bar{p}) + n(1 - k\alpha(\bar{p}))) \).
For one-parameter approximation, we fixed αq = n and

\[ p = \frac{a(\tilde{p})}{(1 - ka(\tilde{p}))}. \] (37)

**Theorem 4.1.** Let \( \hat{T} \) be defined in (34), then

\[ d_{TV}(\hat{T}, Z) \leq (1 - ka(\tilde{p})) \sum_{l=1}^{\infty} |b_l, \tilde{p} - q_{b_l-1, \tilde{p}}| + k(k - 1)a(\tilde{p}) + ka(\tilde{p}) \left( \sum_{l=1}^{\infty} |b_{l, \tilde{p}} - q_{b_{l-1, \tilde{p}}}| + 1 \right), \] (38)

where \( Z \sim \text{NB}(\alpha, p) \), and \( b_{l, \tilde{p}} \) are defined in (36).

**Remarks 4.1.**

(i) The bound in Theorem 4.1 is of constant order and can be calculated for different values of \((k_1, k_2)\) and \(\tilde{p}\).

(ii) In Table 1, the bound for various values of \((k_1, k_2)\) and \(\tilde{p}\) is calculated by taking \(l\) up to the first 3000 terms and neglecting the remainder, as the values are too small. Also, we can observe the pattern that, as the value of \(\tilde{p}\) decreases the bound decreases, which is consistent with NB convergence to Poisson.

### 4.2. Two-parameter approximation

Next, we derive the bound between \( \hat{T} \) and \( Z \) by matching mean and variance as

\[ \frac{\alpha q}{p} = \frac{n(1 - ka(\tilde{p}))}{a(\tilde{p})} \quad \text{and} \quad \frac{\alpha q}{p^2} = n \left( \frac{1 - (2k - 1)a(\tilde{p})}{a(\tilde{p})^2} \right). \] (39)

This leads to the following choice of parameters:

\[ p = \frac{(1 - ka(\tilde{p}))a(\tilde{p})}{1 - (2k - 1)a(\tilde{p})} \quad \text{and} \quad \alpha = \frac{n(1 - ka(\tilde{p}))^2}{1 - 2ka(\tilde{p}) + ka(\tilde{p})^2}. \] (40)
Table 2. Two-parameter approximation.

| $k_1, k_2$ | $p = 1/4$ | $p = 1/8$ | $p = 1/16$ |
|-----------|---------|---------|---------|
|           | $n = 50$ | $n = 100$ | $n = 50$ | $n = 100$ | $n = 50$ | $n = 100$ |
| (1,4)     | 1.1293  | 0.799532| 0.0318133| 0.0225235| 0.0000787781| 0.0000557739|
| (1,5)     | 0.645954| 0.457328| 0.00037684| 0.000266798| 8.05622×10^{-6}| 5.70371×10^{-6}|
| (1,6)     | 0.046521| 0.0329363| 0.0000529935| 0.0000375817| 8.80547×10^{-7}| 6.23417×10^{-7}|
| (1,7)     | 0.00238398| 0.00168783| 0.000105207| 7.44852×10^{-6}| 8.80545×10^{-8}| 6.23415×10^{-8}|
| (1,8)     | 0.000459553| 0.000325358| 1.97243×10^{-7}| 1.39645×10^{-6}| 8.25511×10^{-9}| 5.84452×10^{-9}|
| (1,9)     | 0.000151514| 0.000109819| 3.52218×10^{-7}| 2.49366×10^{-7}| 7.37063×10^{-10}| 5.21832×10^{-10}|
| (2,1)     | 1.64106| 1.16185| 0.0345507| 0.0244615| 0.00013857| 0.000091058|
| (2,5)     | 0.535309| 0.391878| 0.00049724| 0.00035958| 0.000032146| 9.35579×10^{-6}|
| (2,6)     | 0.0305958| 0.0216641| 0.000739922| 0.0000523856| 1.32082×10^{-6}| 9.35125×10^{-7}|
| (2,7)     | 0.00198307| 0.00140399| 0.000138079| 9.77585×10^{-6}| 1.23827×10^{-7} | 8.76677×10^{-8}|
| (2,8)     | 0.000477953| 0.000338385| 2.46553×10^{-6}| 1.74557×10^{-6}| 1.10559×10^{-8} | 7.82478×10^{-9}|
| (3,4)     | 2.09053| 1.48007| 0.0350269| 0.0247986| 0.00021872| 0.000154851|
| (3,5)     | 0.417545| 0.295617| 0.000631458| 0.000447065| 0.000198194| 0.000140319|
| (3,6)     | 0.0197698| 0.0139968| 0.000969564| 0.000668644| 1.85746×10^{-6} | 1.31502×10^{-6}|
| (3,7)     | 0.0007595| 0.00024571| 0.000172595| 0.000122196| 6.58393×10^{-7} | 1.7412×10^{-7}|
| (4,4)     | 2.30166| 1.62955| 0.0339291| 0.0240214| 0.000319874| 0.000226467|
| (4,5)     | 0.286267| 0.202673| 0.0007771| 0.0005518| 0.0000278687| 0.0000197307|
| (4,6)     | 0.013037| 0.00923002| 0.000121071| 0.0000857165| 2.48759×10^{-6} | 1.76119×10^{-6}|
| (5,4)     | 2.18446| 1.54657| 0.0319059| 0.0223589| 0.00044202| 0.000312945|
| (5,5)     | 0.183437| 0.129871| 0.000934084| 0.000658718| 0.000037219| 0.0000264235|
| (6,4)     | 1.80936| 1.281| 0.0294761| 0.0206687| 0.000584608| 0.000413895|

Theorem 4.2. Let $\hat{T}$ be defined in (34) with $1 - 2ka(\hat{p}) + ka(\hat{p})^2 > 0$, then

$$d_{TV}(\hat{T}, Z) \leq \frac{n}{\alpha q} \sqrt{\frac{2}{\pi}} \left( \frac{1}{4} + n \left( 1 - \frac{a(\hat{p})}{2} \left( 1 + a(\hat{p}) \right) \right) \right)^{-1/2} \left[ \sum_{l=1}^{\infty} \frac{l(l-1)}{2} \left| b_{l,\hat{p}} - q b_{l-1,\hat{p}} \right| \right. + \left. \frac{k(k-1)(k-2)}{2} a(\hat{p}) + ka(\hat{p}) \sum_{l=k}^{\infty} \frac{l(l-1)}{2} \left| b_{l-k+1,\hat{p}} - q b_{l-k,\hat{p}} \right| \right],$$

(41)

where $Z \sim \text{NB}(\alpha, p)$ and $b_{l,\hat{p}}$ are defined in (36).

Remarks 4.2.

(i) The bound in Theorem 4.2 is of the order $O(n^{-1/2})$. Therefore, as $n$ increases the bound decreases.

(ii) It is easy to see that the bound in two-parameter approximation is better than one-parameter approximation (see Tables 1 and 2), as expected.

(iii) In Table 2, the bound for various values of $(k_1, k_2)$ and $\hat{p}$ is calculated by taking $l$ up to the first 3000 terms and neglecting the remainder, as the values are too small. Also, we can observe the pattern that, as the value of $\hat{p}$ decreases the bound decreases, which is consistent with NB convergence to Poisson.

4.3. Proofs

Proof of Theorem 4.1. Differentiating (35) w.r.t. $z$, for $|z - a(\hat{p})z^k| < 1$, we get

$$M'_T(z) = n M'_T(z) \left( \frac{1 - ka(\hat{p})z^{k-1}}{1 - z + a(\hat{p})z^k} \right) = n M'_T(z) (1 - ka(\hat{p})z^{k-1})$$

$$\times \sum_{m=0}^{\infty} \left( \sum_{l=0}^{m/k} (-1)^l \binom{m-l(k-1)}{l} a(\hat{p})^l \right) z^m$$
where \( \tilde{p}_m = \mathbf{P}(\hat{T} = m) \) and \( b_{m,\tilde{p}} \) are defined in (36). Multiplying by \((1 - qz)\) and collecting the coefficients of \( z^m \), we get the recurrence relation

\[
q \left( \frac{n}{q} + m \right) \tilde{p}_m - (m + 1) \tilde{p}_{m+1} + n \sum_{l=0}^{m} \tilde{p}_l (b_{m-l,\tilde{p}} - q b_{m-l-1,\tilde{p}}) - ka(\tilde{p}) \sum_{l=0}^{m-k} \tilde{p}_l b_{m-k-l,\tilde{p}} + q k a(\tilde{p}) \sum_{l=0}^{m-k} \tilde{p}_l b_{m-k-l,\tilde{p}} = 0,
\]

where \( q = 1 - p \in (0, 1) \) is defined in (37). Let \( g \in \mathcal{G}_{\hat{T}} \), defined in (5), then

\[
\sum_{m=0}^{\infty} g(m + 1) \left\{ q \left( \frac{n}{q} + m \right) \tilde{p}_m - (m + 1) \tilde{p}_{m+1} + n \sum_{l=0}^{m} \tilde{p}_l (b_{m-l,\tilde{p}} - q b_{m-l-1,\tilde{p}}) - n k a(\tilde{p}) \sum_{l=0}^{m-k} \tilde{p}_l b_{m-k-l,\tilde{p}} + n q k a(\tilde{p}) \sum_{l=0}^{m-k} \tilde{p}_l b_{m-k-l,\tilde{p}} \right\} \tilde{p}_m = 0.
\]

This leads to the following:

\[
\sum_{m=0}^{\infty} \left[ q \left( \frac{n}{q} + m \right) g(m + 1) - mg(m) + n \sum_{l=1}^{\infty} g(m + l + 1) (b_{l,\tilde{p}} - q b_{l-1,\tilde{p}}) - n k a(\tilde{p}) \sum_{l=1}^{\infty} g(m + l + 1) b_{l-1,\tilde{p}} + n q k a(\tilde{p}) \sum_{l=1}^{\infty} g(m + l + 1) b_{l-1,\tilde{p}} \right] \tilde{p}_m = 0.
\]

Hence, the Stein operator of \( \hat{T} \) is given by

\[
\mathcal{A}_{\hat{T}} g(m) = q \left( \frac{n}{q} + m \right) g(m + 1) - mg(m) + n \sum_{l=1}^{\infty} g(m + l + 1) (b_{l,\tilde{p}} - q b_{l-1,\tilde{p}}) - n k a(\tilde{p}) \sum_{l=1}^{\infty} g(m + l + 1) b_{l-1,\tilde{p}} + n q k a(\tilde{p}) \sum_{l=1}^{\infty} g(m + l + 1) b_{l-1,\tilde{p}}
\]

\[
= \tilde{A}_Z g(m) + \mathcal{U}_{\hat{T}} g(m),
\]

where \( \tilde{A}_Z \) denote the Stein operator for NB\( \left( \frac{n}{q}, p \right) \). This is a Stein operator for \( \hat{T} \) in perturbation of the NB operator. Using (14) in perturbed operator \( \mathcal{U}_{\hat{T}} \), we get

\[
\mathcal{U}_{\hat{T}} g(m) = n \left[ \sum_{l=1}^{\infty} b_{l,\tilde{p}} - q \sum_{l=1}^{\infty} b_{l-1,\tilde{p}} - k a(\tilde{p}) \sum_{l=1}^{\infty} b_{l-k+1,\tilde{p}} + q k a(\tilde{p}) \sum_{l=1}^{\infty} b_{l-k,\tilde{p}} \right] g(m + 1)
\]

\[
+ n \sum_{l=1}^{\infty} \sum_{j=1}^{l} \Delta g(m + j) b_{l,\tilde{p}} - n q \sum_{l=1}^{\infty} \sum_{j=1}^{l} \Delta g(m + j) b_{l-1,\tilde{p}}
\]

where

\[
\Delta g(m + j) = g(m + j) - g(m + j - 1).
\]
\[ -nka(\tilde{p}) \sum_{l=k-1}^{\infty} \sum_{j=1}^{l} \Delta g(m+j)b_{l-k+1,\tilde{p}} + nqka(\tilde{p}) \sum_{l=k}^{\infty} \sum_{j=1}^{l} \Delta g(m+j)b_{l-k,\tilde{p}} \]

Observe that \( \sum_{l=0}^{\infty} b_{l,\tilde{p}} = 1/a(\tilde{p}) \) and \( b_{0,\tilde{p}} = 1 \). Using (37), we obtain the perturbation operator

\[ \mathcal{U}_fg(m) = n \sum_{l=1}^{\infty} \sum_{j=1}^{l} \Delta g(m+j)b_{l,\tilde{p}} - nq \sum_{l=1}^{\infty} \sum_{j=1}^{l} \Delta g(m+j)b_{l-1,\tilde{p}} - nka(\tilde{p}) \]

\[ \times \sum_{l=k-1}^{\infty} \sum_{j=1}^{l} \Delta g(m+j)b_{l-k+1,\tilde{p}} + nqka(\tilde{p}) \sum_{l=k}^{\infty} \sum_{j=1}^{l} \Delta g(m+j)b_{l-k,\tilde{p}} \]

Hence, for \( g \in \mathcal{G}_Z \cap \mathcal{G}_f \), taking expectation w.r.t. \( \hat{T} \) and using (10), we get required result. \( \square \)

**Proof of Theorem 4.2.** Next, for \( g \in \mathcal{G}_f \), we introduce a new parameter \( \alpha > 0 \) in (42) as

\[ A_\tilde{f}g(m) = q(\alpha + m)g(m+1) - mg(m) + (n - \alpha q)g(m+1) \]

\[ + n \sum_{l=1}^{\infty} g(m+l+1)(b_{l,\tilde{p}} - qb_{l-1,\tilde{p}}) \]

\[ - nka(\tilde{p}) \sum_{l=k-1}^{\infty} g(m+l+1)b_{l-k+1,\tilde{p}} + nqka(\tilde{p}) \sum_{l=k}^{\infty} g(m+l+1)b_{l-k,\tilde{p}} \]

\[ = A_2g(m) + \dot{U}_fg(m) \]

where \( \alpha \) and \( q = 1 - p \) are defined in (40). This is the Stein operator for \( \tilde{T} \), which is a perturbation of \( \text{NB}(\alpha, p) \). Putting (14) and (20) in perturbed operator \( \dot{U}_f \), then using (39), we get

\[ \dot{U}_fg(m) = n \left[ \sum_{l=k-1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta^2 g(m+u)b_{l,\tilde{p}} - q \sum_{l=1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta^2 g(m+u)b_{l-1,\tilde{p}} \right] \]

\[ - ka(\tilde{p}) \sum_{l=k-1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta^2 g(m+u)b_{l-k+1,\tilde{p}} + kqa(\tilde{p}) \sum_{l=k}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta^2 g(m+u)b_{l-k,\tilde{p}} \] \]

Taking the expectation w.r.t. \( \hat{T} \), we have

\[ \mathbb{E} [\dot{U}_fg(\hat{T})] \]

\[ = n \sum_{m=0}^{\infty} \left[ \sum_{l=1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta^2 g(m+u)b_{l,\tilde{p}} - q \sum_{l=1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta^2 g(m+u)b_{l-1,\tilde{p}} \right] \hat{p}_m \]

\[ - ka(\tilde{p}) \sum_{l=k-1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta g(m+u)b_{l-k+1,\tilde{p}} + kqa(\tilde{p}) \sum_{l=k}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta g(m+u)b_{l-k,\tilde{p}} \] \]

\[ = n \sum_{m=0}^{\infty} \left[ \sum_{l=1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta g(m+u)b_{l,\tilde{p}} - q \sum_{l=1}^{\infty} \sum_{j=1}^{l} \sum_{u=1}^{j-1} \Delta g(m+u)b_{l-1,\tilde{p}} \right] (\hat{p}_{m-1} - \hat{p}_m) \]

Now, \( g \in \mathcal{G}_Z \cap \mathcal{G}_f \), taking supremum, and using (10), we get required result. \( \square \)
References

Balakrishnan, N., Koutras, M.V. (2002). *Runs and Scans with Applications*. New York: John Wiley.

Barbour, A.D. (1990). Stein’s method for diffusion approximations. *Probab. Theory Related Fields* 84:297–322.

Barbour, A.D., Čekanavičius, V. (2002). Total variation asymptotes for sum of independent integer random variables. *Ann. Prob.* 30:509–545.

Barbour, A.D., Chen, L.H.Y. (2014). Stein’s (magic) method. Preprint: arXiv:1411.1179.

Barbour, A.D., Chen, L.H.Y., Loh, W.L. (1992a). Compound Poisson approximation for nonnegative random variables via Stein’s method. *Ann. Prob.* 20:1843–1866.

Barbour, A.D., Čekanavičius, V., Xia, A. (2007). On Stein’s method and perturbations. *ALEA* 3:31–53.

Barbour, A.D., Xia, A. (1999). Poisson perturbation. *ESAIM Probab. Stat.* 3:131–150.

Brown, T.C., Phillips, M.J. (1999). Negative binomial approximation with Stein’s method. *Methodol. Comput. Appl. Probab.* 1:407–421.

Chen, L.H.Y., Goldstein, L., Shao, Q.-M. (2011). *Normal Approximation by Stein’s Method*. Heidelberg: Springer.

Dafnis, S.D., Antzoulakos, D.L., Philippou, A.N. (2010). Distribution related to \((k_1, k_2)\) events. *J. Stat. Plan. Inference* 140:1691–1700.

Daly, F. (2010). Stein’s method for compound geometric approximation. *J. Appl. Probab.* 47:146–156.

Daly, F. (2011). On Stein’s method, smoothing estimates in total variation distance and mixture distributions. *J. Stat. Plan. Inference* 141:2228–2237.

Daly, F., Lefevre, C., Utev, S. (2012). Stein’s method and stochastic orderings. *Adv. Appl. Probab.* 44:343–372.

Diaconis, P., Zabell, S. (1991). Closed form summation for classical distributions. Variations on a theme of de Moivre. *Stat. Sci.* 6:284–302.

Goldstein, L., Reinert, G. (2005). Distributional transformations, orthogonal polynomials and Stein characterizations. *J. Theoret. Probab.* 18:237–260.

Götze, F. (1991). On the rate of convergence in the multivariate CLT. *Ann. Prob.* 19:724–739.

Holmes, S. (2004). Stein’s method for birth and death chains. *Inst. Math. Stat. (IMS) Collect.* 46:42–65.

Huang, W.T., Tsai, C.S. (1991). On a modified binomial distribution of order \(k\). *Stat. Probab. Lett.* 11:125–131.

Ley, C., Swan, Y. (2013a). Stein’s density approach and information inequalities. *Electron. Commun. Probab.* 18:1–14.

Ley, C., Swan, Y. (2013b). Local Pinsker inequalities via Stein’s discrete density approach. *IEEE Trans. Inf. Theory* 59:5584–5591.

Ley, C., Reinert, G., Swan, Y. (2017). Stein’s method for comparison of univariate distributions. *Probability Surveys* 14:1–52.

Mattner, L., Roos, B. (2007). A shorter proof of Kanter’s Bessel function concentration bound. *Probab. Theory Relat. Fields* 139:407–421.

Norudin, I., Peccati, G. (2012). Normal approximations with Malliavin calculus. From Stein’s method to universality. *Cambridge Tracts in Mathematics* No. 192.

Reinert, G. (2005). Three general approaches to Stein’s method. An introduction to Stein’s method. *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* (4, 183221). Singapore: Singapore University Press.

Roos, B. (2003). Kerstan’s method for compound Poisson approximation. *Ann. Prob.* 31:1754–1771.

Ross, N. (2011). Fundamentals of Stein’s method. *Probab. Surv.* 8:210–293.

Stein, C. (1986). *Approximate Computation of Expectations*. *Institute of Mathematical Statistics Lecture Notes-Monograph Series* (vol. 7). Hayward, CA: Institute of Mathematical Statistics.

Stein, C., Diaconis, P., Holmes, S., Reinert, G. (2004). Use of exchangeable pairs in the analysis of simulations. In: Diaconis, P., Holmes, S., eds. *Stein’s Method: Expository Lectures and Applications*. *IMS Lecture Notes Monogr. Ser* 46 1–26. Beachwood, Ohio, USA: Institute of Mathematical Statistics.
Upadhye, N.S., Čekanavičius, V., Vellaisamy, P. (2014). On Stein operators for discrete approximations. *Bernoulli* 23:2828–2859.

Vellaisamy, P., Upadhye, N.S., Čekanavičius, V. (2013). On negative binomial approximation. *Theory Probab. Appl.* 57(1):97–109.

Vellaisamy, P., Upadhye, N.S. (2003). Compound negative binomial approximations for sums of random variables. *Probab. Math. Stat.* 31:205–226.

Wang, X., Xia, A. (2008). On negative approximation to $k$-runs. *J. Appl. Probab.* 45:456–471.

Yakshyavichus, S.H. (1998). On a method of expansion of the probabilities of lattice random variables. *Theory Probab. Appl.* 42(2):271–282.