PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 29, Fasc. 2 (2009), pp. 205–226

COMPOUND NEGATIVE BINOMIAL APPROXIMATIONS FOR SUMS OF RANDOM VARIABLES

BY

P. VELLAISAMY AND N. S. UPADHYE* (MUMBAI)

Abstract. The negative binomial approximations arise in telecommunications, network analysis and population genetics, while compound negative binomial approximations arise, for example, in insurance mathematics. In this paper, we first discuss the approximation of the sum of independent, but not identically distributed, geometric (negative binomial) random variables by a negative binomial distribution, using Kerstan's method and the method of exponents. The appropriate choices of the parameters of the approximating distributions are also suggested. The rates of convergence obtained here improve upon, under certain conditions, some of the known results in the literature. The related Poisson convergence result is also studied. We then extend Kerstan's method to the case of compound negative binomial approximations and error bounds for the total variation metric are obtained. The approximation by a suitable finite signed measure is also studied. Some interesting special cases are investigated in detail and a few examples are discussed as well.

2000 AMS Mathematics Subject Classification: Primary: 60E05; Secondary: 62E17, 60F05.

Key words and phrases: Compound negative binomial approximation, compound Poisson approximation, finite signed measure, Kerstan's method, method of exponents, total variation distance.

1. INTRODUCTION

Let $\{X_i\}$ be a sequence of discrete random variables (rv's) and $S_n = \sum_{j=1}^{n} X_j$. When the X_i are independent $B(p_i)$ variables, it is known that (see Khintchine [15] or Le Cam [16])

$$d_{TV}\big(S_n, P(\lambda)\big) \leqslant \sum_{i=1}^{n} p_i^2, \tag{1.1}$$

* The author thanks CSIR for the award of a research fellowship.

where, for any two rv's X and Y,

$$d_{TV}(X,Y) = \sup_A |P(X \in A) - P(Y \in A)|$$

denotes the total variation metric and $P(\lambda)$ denotes the Poisson variable with parameter $\lambda = \sum_{i=1}^{n} p_i$. Kerstan [14] improved the above bound to

$$d_{TV}\big(S_n, P(\lambda)\big) \leqslant 1.05\big(\sum_{i=1}^{n} p_i^2\big)/\big(\sum_{i=1}^{n} p_i\big)$$

when $p_i \leqslant 1/4$. Barbour and Hall [3] further improved the bound to

$$d_{TV}\big(S_n, P(\lambda)\big) \leqslant \sum_{i=1}^{n} p_i^2 \min\{1, \lambda^{-1}\},$$

where λ^{-1} is known as the 'magic factor'. See, also Barbour et al. [4] for more details and related results. The negative binomial (NB) approximation to the sum of indicator rv's is studied by Brown and Phillips [5]. They showed that NB distribution arises also as the limiting distribution of Pólya distribution. Also, Brown and Xia [6] considered the NB approximation to the number of two-runs.

In this paper, we first consider the problem of approximating the distribution of $S_n = \sum_{i=1}^{n} X_i$, where X_i's are independent geometric $Ge(p_i)$ variables, by an $NB(r,p)$ distribution. In Section 2, we obtain the upper bounds similar to (1.1), using Kerstan's method (Kerstan [14]) and generalize the results to the sum of independent negative binomial random variables using the method of exponents (Čekanavičius and Roos [8]–[10]). Also, the conditions under which $S_n \xrightarrow{\mathcal{L}} P(\lambda)$ are investigated. Kerstan's method is originally due to Kerstan [14], which was later modified and used by several authors, say, for example, Daley and Vere-Jones ([11], pp. 187–190), Witte [30] and Roos [19], [20], where the problems of Poisson, multivariate Poisson and compound Poisson (CP) approximations are considered. Recently, Barbour [2] obtained a bound, using Stein's method, for multivariate Poisson-binomial distribution. This bound is comparable to the one obtained by Roos [19] using Kerstan's method. Also, Roos [20] studied CP approximations for the sums of independent discrete valued rv's using Kerstan's method. Čekanavičius [7] considered approximation of compound distributions using Le Cam's [16] operator theoretic method.

Note that the CP distribution plays an important role in risk theory. Consider the model for the total claim amount in a portfolio of insurance policies. Let N denote the number of claims arising from policies in a given period of time, and let X_i denote the claim size amount of the i-th claim. In a collective risk model, the random sum $S_N = \sum_{i=1}^{N} X_i$ denotes the aggregate claim, where it is assumed that X_i's are iid and N is independent of the X_i. In many applications, the number of claims N is assumed to follow Poisson distribution, and thus the distribution of the random sum S_N is a CP distribution. However, there are many situations (see, for example, Panjer and Willmot [18], Drekic and Willmot [12]) and, in particular, in non-life insurance modeling, the CP distribution may not serve as a suitable model for the insurance data. In such cases, and especially when $\mathbb{V}(N) > \mathbb{E}(N)$, an NB model is usually suggested. Thus, the study of the compound negative binomial (CNB) distribution arises naturally, and hence the CNB approximation problems are of importance. In Section 3, we use Kerstan's method to obtain the first-order result for the total variation distance. Also, we study the approximation of S_n by a finite signed measure, which leads to the improvements in the constant that appears in the first-order result. In Section 4, we apply our results to the distributions which can be written as infinite mixtures and obtain results analogous to the ones given in Section 3. Finally, an application of our results to life and health insurance is also pointed out.

2. NEGATIVE BINOMIAL AND POISSON APPROXIMATIONS

Throughout the paper, $\mathbb{Z}_+ = \{0, 1, \ldots\}$ denotes the set of non-negative integers. Let $X_i \sim Ge(p_i)$, $1 \leqslant i \leqslant n$, be independent geometric rv's with

$$P(X_i = k) = p_i(1-p_i)^k \quad \text{for } k \in \mathbb{Z}_+, \; 0 < p_i < 1.$$

We are interested in approximating the distribution of the sum $S_n = \sum_{j=1}^{n} X_j$ to an $NB(r, p)$ rv N with

$$P(N = k) = \binom{r+k-1}{k} p^r (1-p)^k \quad \text{for } k \in \mathbb{Z}_+,$$

where $r > 0$ and $0 < p < 1$, and also choosing the appropriate parameters r and p so that the error is minimum. We adopt Kerstan's method (see Roos [19]) and the method of exponents (see Čekanavičius and Roos [8]–[10]) for finding upper bounds for $d_{TV}(S_n, N)$.

2.1. Kerstan's method. For the power series $f(z) = \sum_{m=0}^{\infty} a_m z^m$, where $a_m \in \mathbb{R}$ and $z \in \mathbb{C}$, define the norm $\|f(z)\| = \sum_{m=0}^{\infty} |a_m|$. Then, it is well known that $\|f_1(z) f_2(z)\| \leqslant \|f_1(z)\| \|f_2(z)\|$. For a $\mathbb{Z}_+$-valued rv X, the probability generating function (pgf)

$$\mathbb{E}(z^X) = \sum_{m=0}^{\infty} z^m P(X = m),$$

where $z \in \mathbb{C}$ and $|z| \leqslant 1$, exists. Then, for $X \geqslant 0$ and $Y \geqslant 0$ (see Roos [19]),

$$d_{TV}(X, Y) = \frac{1}{2} \|\mathbb{E}(z^X) - \mathbb{E}(z^Y)\|. \tag{2.1}$$

We now obtain the norm estimates for $\gamma(z) = \mathbb{E}(z^{S_n}) - \mathbb{E}(z^N)$ for the case $r = n$.

Note that

$$\gamma(z) = \prod_{i=1}^{n}\left(\frac{p_i}{1-q_iz}\right) - \left(\frac{p}{1-qz}\right)^n \tag{2.2}$$

$$= \Big[\prod_{i=1}^{n}(1+L_i) - 1\Big]\left(\frac{p}{1-qz}\right)^n \tag{2.3}$$

$$= \sum_{j=1}^{n}\sum_{1\leqslant i_1<\ldots<i_j\leqslant n}\prod_{s=1}^{j}\left[L_{i_s}\left(\frac{p}{1-qz}\right)^{n/j}\right], \tag{2.4}$$

where

$$L_i = \frac{p_i(1-qz)}{p(1-q_iz)} - 1 = \left(1-\frac{p_i}{p}\right)\frac{(z-1)}{(1-q_iz)}.$$

Therefore, using Lemma 3.1 of Čekanavičius and Roos [8] and $\|1/(1-q_iz)\| = 1/p_i$, we get

$$\left\|L_i\left(\frac{p}{1-qz}\right)^{n/j}\right\| \leqslant \frac{1}{p_i}\left|1-\frac{p_i}{p}\right|\left\|(z-1)\exp\left(\frac{n}{j}\sum_{m=1}^{\infty}\frac{q^m}{m}(z^m-1)\right)\right\|$$

$$\leqslant \frac{1}{p_i}\left|1-\frac{p_i}{p}\right|\big\|(z-1)\exp\big((nq/j)(z-1)\big)\big\| \leqslant \frac{1}{p_i}\left|1-\frac{p_i}{p}\right|\min\left\{2,\sqrt{\frac{2j}{nqe}}\right\},$$

since

$$\|h(z)\| = \big\|\exp\big\{(n/j)\sum_{m=2}^{\infty}q^m(z^m-1)/m\big\}\big\| = 1.$$

For, if $\lambda_m = nq^m/(jm)$, $\lambda = \sum_{m=2}^{\infty}\lambda_m$ and $Q(\{m\}) = \lambda_m/\lambda$, then $h(z) = \mathbb{E}(z^X)$, where $X \sim CP(\lambda, Q)$. Hence, from (2.4) we obtain

$$\begin{aligned}\|\gamma(z)\| &\leqslant \sum_{j=1}^{n}\sum_{1\leqslant i_1<i_2<\ldots<i_j\leqslant n}\prod_{s=1}^{j}\left\|L_{i_s}\left(\frac{p}{1-qz}\right)^{n/j}\right\| \\ &\leqslant \sum_{j=1}^{n}\frac{1}{j!}\left[\sum_{i=1}^{n}\left\|L_i\left(\frac{p}{1-qz}\right)^{n/j}\right\|\right]^j \\ &\leqslant \sum_{j=1}^{n}\frac{1}{j!}\left[\sum_{i=1}^{n}\frac{1}{p_i}\left|1-\frac{p_i}{p}\right|\min\left\{2,\sqrt{\frac{2j}{nqe}}\right\}\right]^j \\ &\leqslant \sum_{j=1}^{n}\frac{(\sqrt{j}\alpha_n)^j}{j!},\end{aligned} \tag{2.5}$$

where

$$\alpha_n = \sum_{i=1}^{n}\frac{1}{p_i}\left|1-\frac{p_i}{p}\right|\min\left\{2,\sqrt{\frac{2}{nqe}}\right\}. \tag{2.6}$$

Thus, from (2.1)–(2.5), we have the following result.

THEOREM 2.1. *Let X_i, $1 \leqslant i \leqslant n$, be a sequence of independent geometric $Ge(p_i)$ variables*, $S_n = \sum_{j=1}^{n} X_j$, *and* $N \sim NB(n,p)$. *Then*

$$d_{TV}(S_n, N) \leqslant \min\{1.37\alpha_n, 1\}, \tag{2.7}$$

and, *when* $\alpha_n < e$,

$$d_{TV}(S_n, N) \leqslant \frac{\alpha_n/e}{2\sqrt{2\pi}(1-\alpha_n/e)}, \tag{2.8}$$

where α_n is given by (2.6).

The first estimate (referred to as a practical estimate), given in (2.7), follows from the fact that

$$d_{TV}(S_n, N) \leqslant \min\{f(\alpha_n), 1\} \leqslant \min\left\{\frac{\alpha_n}{x_0}, 1\right\},$$

where $f(x) = \frac{1}{2}\sum_{j=1}^{\infty}(x\sqrt{j})^j/j!$ and, numerically, it can be seen that $0.73 < x_0 < 0.74$ is the unique solution of $f(x) = 1$. The later estimate follows by an application of Stirling's approximation formula and is less than one when $\alpha_n < 2.26$.

REMARK 2.1. (i) If $p_i = p$, then it follows from (2.7) that $d_{TV}(S_n, N) = 0$, which holds iff $S_n \sim NB(n,p)$, as expected.

(ii) Observe that the bound given in (2.7) contains a 'magic factor' $(nq)^{-1/2}$ which improves the estimate for large n.

(iii) Let $\mu_n = \sum_{i=1}^{n}(q_i/p_i)$. One way to choose p is such that $\mathbb{E}(S_n) = \mathbb{E}(N)$, which leads to the choice $p = n/(n+\mu_n)$ and the upper bound for this case can be obtained from the practical estimate given in (2.7). The other choices of p for better accuracy are $p = \min_i p_i$ and $p = \max_i p_i$.

Next, we obtain some improvements over the above results along with generalizations for the sum of independent NB variables.

2.2. The method of exponents. In this subsection, we obtain an $NB(r,p)$ approximation result for $S_n = \sum_{j=1}^{n} X_j$, where $X_j \sim NB(\alpha_j, p_j)$, using the method of exponents. Here, we assume that $r = \alpha = \sum_{i=1}^{n} \alpha_i$ and $p = \big(\sum_{i=1}^{n} \alpha_i p_i\big)/\alpha$ and write the pgf of S_n as

$$\begin{aligned} \mathbb{E}(z^{S_n}) &= \prod_{j=1}^{n}\left(\frac{p_j}{1-q_j z}\right)^{\alpha_j} = \exp\left(\sum_{j=1}^{n}\alpha_j \ln\left(\frac{p_j}{1-q_j z}\right)\right) \\ &= \exp\left(\sum_{m=1}^{\infty}\frac{1}{m}\sum_{j=1}^{n}\alpha_j q_j^m (z^m - 1)\right) := \exp(F). \end{aligned} \tag{2.9}$$

Similarly,

$$\mathbb{E}(z^N) = \exp\left(\alpha \sum_{m=1}^{\infty} \frac{q^m}{m}(z^m - 1)\right) := \exp(A). \tag{2.10}$$

Now

$$\begin{aligned}\|\exp(F) - \exp(A)\| &= \left\|\exp(A) \int_0^1 \Big(\exp\big(x(F-A)\big)\Big)' dx\right\| \\ &\leqslant \int_0^1 \big\|(F-A)\exp\big(xF + (1-x)A\big)\big\| dx,\end{aligned}$$

where the prime ($'$) denotes the derivative with respect to x. Substituting the power series expansion for F and A from (2.9) and (2.10), respectively, we obtain, for $0 < x < 1$,

$$xF + (1-x)A = \alpha q(z-1) + T,$$

where

$$T = \sum_{m=2}^{\infty} \frac{1}{m}\big(\sum_{j=1}^{n} x\alpha_j q_j^m + (1-x)\alpha q^m\big)(z^m - 1).$$

Also, observe that $\exp(T)$ forms a compound distribution, as all the multipliers of $(z^m - 1)$ in T are non-negative, and hence $\|\exp(T)\| = 1$. Therefore,

$$\begin{aligned}\big\|\exp(F) - \exp(A)\big\| &\leqslant \big\|(F-A)\exp\big(\alpha q(z-1)\big)\big\| \\ &\leqslant \sum_{m=2}^{\infty} \frac{1}{m}\big(\sum_{j=1}^{n} \alpha_j q_j^m - \alpha q^m\big)\big\|(z^m - 1)\exp\big(\alpha q(z-1)\big)\big\| \\ &\leqslant \left(\sum_{j=1}^{n} \frac{\alpha_j q_j^2}{p_j} - \frac{\alpha q^2}{p}\right)\big\|(z-1)\exp\big(\alpha q(z-1)\big)\big\| \\ &\leqslant \left(\sum_{j=1}^{n} \frac{\alpha_j q_j^2}{p_j} - \frac{\alpha q^2}{p}\right)\min\left\{2, \sqrt{\frac{2}{\alpha q e}}\right\},\end{aligned}$$

where the non-negativity in the second inequality is implied by the fact that $\alpha q^m \leqslant \sum_{j=1}^{n} \alpha_j q_j^m$, which in turn follows from a simple application of Jensen's inequality. Also, the last inequality follows from Lemma 3.1 of Čekanavičius and Roos [8]. Thus, we obtain the following result from (2.1).

THEOREM 2.2. *For $\alpha_i > 0$, let $X_i \sim NB(\alpha_i, p_i)$, $1 \leqslant i \leqslant n$, be a sequence of independent random variables, $S_n = \sum_{i=1}^{n} X_i$, and $N \sim NB(\alpha, p)$, where $\alpha = \sum_{i=1}^{n} \alpha_i$ and $p = \sum_{i=1}^{n} \alpha_i p_i / \alpha$. Then*

$$d_{TV}(S_n, N) \leqslant \left(\sum_{j=1}^{n} \frac{\alpha_j q_j^2}{p_j} - \frac{\alpha q^2}{p}\right)\min\left\{1, \frac{1}{\sqrt{2\alpha q e}}\right\}. \tag{2.11}$$

REMARK 2.2. Observe that the bound in (2.11), obtained using the method of exponents, also involves the 'magic factor' $(\alpha q)^{-1/2}$ and is comparable to the one given in (2.7) obtained by Kerstan's method.

2.3. NB to Poisson approximation. Let $N \sim NB(r, p)$, where $r > 0$ and $0 < p < 1$, and $Y \sim P(\lambda)$. Here, we follow the approach of expansion in the exponents. Write the pgf of N (see (2.10)) as

$$\mathbb{E}(z^N) = \exp\left(r \sum_{m=1}^{\infty} \frac{q^m}{m}(z^m - 1)\right) = \exp(A).$$

Also, the pgf of Y is $\exp\big(\lambda(z-1)\big) := \exp(B)$. Then, as seen earlier,

$$\|\exp(A) - \exp(B)\| \leqslant \int_0^1 \big\|(A - B)\exp\big(xA + (1-x)B\big)\big\| dx.$$

Moreover,

$$xA + (1-x)B = rq(z-1) + \sum_{m=2}^{\infty} \frac{1}{m}\big(xq^m(z^m - 1)\big) := rq(z-1) + M.$$

Then $\|\exp(M)\| = 1$, as all the multipliers of $(z^m - 1)$ are non-negative. Letting $rq = \lambda$ and following the arguments similar to the derivation of Theorem 2.2, we obtain

THEOREM 2.3. *Let $N \sim NB(r, p)$ and $Y \sim P(\lambda)$, where $\lambda = rq$ and $q = 1 - p$. Then*

$$d_{TV}(N, Y) \leqslant \frac{rq^2}{p} \min\left\{1, \frac{1}{\sqrt{2rqe}}\right\}. \tag{2.12}$$

REMARK 2.3. (i) The bound given above contains the 'magic factor' $\lambda^{-1/2}$ which reduces the bound considerably when $\lambda = rq$ is large.

(ii) The best available bound in the literature for this case is given by Roos [22] (see also Roos [21]), namely,

$$d_{TV}(N, Y) \leqslant \frac{rq}{p^2} \min\left\{\frac{3p}{4rqe}, 1\right\}. \tag{2.13}$$

It can be seen that our bound in (2.12) is better than the above one, when $r \leqslant \big(3/(4q^2 e)\big) \min\{1, 3/(2q)\}$.

Finally, we give below the rate of convergence result for Poisson approximation to S_n, obtained using again the method of exponents.

THEOREM 2.4. *Let* $X_i \sim NB(\alpha_i, p_i)$, $1 \leqslant i \leqslant n$, *be independent random variables*, *and* $S_n = \sum_{i=1}^n X_i$. *Then*

$$d_{TV}\big(S_n, P(\lambda)\big) \leqslant \sum_{j=1}^{n} \frac{\alpha_j q_j^2}{p_j} \min\left\{1, \frac{1}{\sqrt{2\lambda e}}\right\}, \tag{2.14}$$

where $\lambda = \sum_{i=1}^n \alpha_i q_i = \alpha q$.

REMARK 2.4. (i) It is interesting to note that the bound in (2.14) is nothing but the sum of the bounds in (2.11) and (2.12).

(ii) A comparison of bounds (2.11) and (2.14) shows that an NB approximation is better than Poisson approximation in the case of sum of independent $NB(\alpha_i, p_i)$ variables. This motivates our study of approximation by compound distributions and, in particular, to CNB distribution in the next section.

COROLLARY 2.1. *Let* $\{X_{ni}\}$ *be a sequence of* $Ge(p_{ni})$ *variables*, *and* $S_n = \sum_{i=1}^n X_{ni}$. *If* $\max_{1 \leqslant i \leqslant n} q_{ni} \to 0$ *so that* $\sum_{i=1}^n q_{ni} \to \lambda > 0$ *as* $n \to \infty$, *then*

$$S_n \xrightarrow{\mathcal{L}} P(\lambda).$$

The corollary follows from (2.14) and the fact that

$$0 \leqslant \lim_{n\to\infty} \sum_{j=1}^{n} \frac{q_{nj}^2}{p_{nj}} \leqslant \lim_{n\to\infty} \big(\max_{1\leqslant j\leqslant n} q_{nj}\big) \lim_{n\to\infty} \sum_{j=1}^{n} \frac{q_{nj}}{p_{nj}} = 0.$$

The above result is essentially due to Wang [29].

3. APPROXIMATION BY COMPOUND DISTRIBUTIONS

In this section, we study CNB approximation to S_n, where $S_n = \sum_{i=1}^n X_i$, and $X_i \sim F_i$, a discrete real-valued distribution.

3.1. Preliminary results. Let μ be a finite signed measure on $(\mathbb{R}, \mathcal{B})$. A measurable set B is said to be a *positive set with respect to* μ, denoted by $B \geqslant 0$, if $\mu(A \cap B) \geqslant 0$ for every $A \in \mathcal{B}$. Similarly, a set C is called a *negative set with respect to* μ, denoted by $C \leqslant 0$, if $\mu(C \cap A) \leqslant 0$ for every $A \in \mathcal{B}$. Also, a pair (B, C) is said to be the *Hahn decomposition* of $\mathbb{R}$ if $B \cup C = \mathbb{R}$, where $B \geqslant 0$ and $C \leqslant 0$. Also, the total variation norm of μ is defined (see, for example, Aliprantis and Burkinshaw [1]; Rudin [23])) as

$$\begin{aligned} \|\mu\| = |\mu|(\mathbb{R}) &= \mu_+(\mathbb{R}) + \mu_-(\mathbb{R}) = \mu(B) - \mu(C) \\ &= 2\mu(B) - \mu(\mathbb{R}), \end{aligned} \tag{3.1}$$

where μ_+ and μ_- are positive and negative variations of μ. It is well known that the total variation distance d_{TV} between the distribution of two discrete rv's X and Y is

$$d_{TV}(X,Y) = \sup_A |P(X \in A) - P(Y \in A)| \tag{3.2}$$

$$= P(X \in D) - P(Y \in D)$$

$$= (P_X - P_Y)(D) := \mu_{X,Y}(D), \tag{3.3}$$

where $D = \{m : P(X = m) \geqslant P(Y = m)\} = \big\{m : \mu_{X,Y}\{m\} \geqslant 0\big\}$ is a positive set of $\mu_{X,Y}$ (Wang [28] or Vellaisamy and Chaudhuri [25]). Also, by (3.1), we have

$$\|\mu_{X,Y}\| = 2\mu_{X,Y}(D). \tag{3.4}$$

Thus, from (3.3) and (3.4) we obtain the relation

$$d_{TV}(X,Y) = \mu_{X,Y}(D) = \frac{1}{2}\|\mu_{X,Y}\|,$$

where D is a positive set of $\mu_{X,Y}$.

3.2. CNB and CP approximations. Let Y_i be iid real-valued rv's with distribution Q, $N \sim NB(r,p)$, and $M \sim P(\lambda)$, $\lambda > 0$. Assume that the Y_i, N and M are independent. Then the distribution of $T_N = \sum_{j=1}^{N} Y_j$ is the CNB distribution with parameters r, p and Q, and is denoted by $CNB(r,p,Q)$. Since

$$P(T_N \in A) = \sum_{k=0}^{\infty} P(N = k) Q^{*k}(A),$$

we get

$$CNB(r,p,Q) = \left(\frac{p\delta_0}{\delta_0 - qQ}\right)^r, \tag{3.5}$$

where $q = 1 - p$, Q^{*k} denotes the k-fold convolution of Q, and δ_0 is the Dirac measure at 0. Similarly, the distribution of T_M, denoted by $CP(\lambda, Q)$, is the CP distribution with parameters λ and Q, and its distribution is $\exp\big(\lambda(Q - \delta_0)\big)$. First, we obtain the error bounds in the approximation of $CNB(r,p,Q)$ to $CP(\lambda,Q)$.

THEOREM 3.1. *Let $r > 0$, Q be any distribution on $\mathbb{R}$ and $rq = \lambda$. Then*

$$\sup_Q \left\| \left(\frac{p\delta_0}{\delta_0 - qQ}\right)^r - \exp\big(\lambda(Q - \delta_0)\big) \right\| \leqslant \frac{rq^2}{p} \min\left\{2, \sqrt{\frac{2}{\lambda e}}\right\}.$$

Proof. First note that for every Q

$$\begin{aligned}\left\|\left(\frac{p\delta_0}{\delta_0 - qQ}\right)^r - \exp\big(\lambda(Q-\delta_0)\big)\right\| &\\ &\leqslant \sum_{m=0}^{\infty}\left|\binom{r+m-1}{m}p^r q^m - \frac{e^{-\lambda}\lambda^m}{m!}\right| \|Q\|^m \\ &= \left\|\left(\frac{p\delta_0}{\delta_0 - q\delta_1}\right)^r - \exp\big(\lambda(\delta_1-\delta_0)\big)\right\| = 2d_{TV}(N,Y),\end{aligned}$$

and now the result follows from Theorem 2.3. ∎

REMARK 3.1. The above result follows easily also from Theorem 2.3 and Lemma 3.1 of Vellaisamy and Chaudhuri [24].

Next we consider the approximation of finite sum $S_n = \sum_{j=1}^n Z_j$, where $Z_j = \sum_{i=1}^{N_j} X_i$, and $N_j \sim NB(\alpha_j, p_j)$, to the distributions $CNB(n,p,Q)$ and $CP(\lambda, Q)$. Note that $Z_j \sim CNB(\alpha_j, p_j, Q) = \big(p_j\delta_0/(\delta_0 - q_jQ)\big)^{\alpha_j}$, the compound negative binomial distribution with parameters $\alpha_j > 0$, p_j and Q.

THEOREM 3.2. *Let $S_n = \sum_{j=1}^n Z_j$, where Z_j's are independent with distributions $CNB(\alpha_j, p_j, Q)$. Then, for any distribution Q on $\mathbb{R}$, we have*

$$\begin{aligned}\text{(3.6)}\quad \sup_Q \left\|\prod_{i=1}^n \left(\frac{p_i\delta_0}{\delta_0 - q_iQ}\right)^{\alpha_i} - \left(\frac{p\delta_0}{\delta_0 - qQ}\right)^{\alpha}\right\| &\\ &\leqslant \left(\sum_{i=1}^n \frac{\alpha_i q_i^2}{p_i} - \frac{\alpha q^2}{p}\right)\min\left\{2, \sqrt{\frac{2}{\alpha q e}}\right\};\end{aligned}$$

$$\begin{aligned}\text{(3.7)}\quad \sup_Q \left\|\prod_{i=1}^n \left(\frac{p_i\delta_0}{\delta_0 - q_iQ}\right)^{\alpha_i} - \exp\big(\lambda(Q-\delta_0)\big)\right\| &\\ &\leqslant \sum_{i=1}^n \frac{\alpha_i q_i^2}{p_i}\min\left\{2, \sqrt{\frac{2}{\lambda e}}\right\},\end{aligned}$$

where $\alpha = \sum_{i=1}^n \alpha_i$, $p = \sum_{i=1}^n \alpha_i p_i/\alpha$, $q = 1-p$ and $\lambda = \alpha q$.

Proof. The bound in (3.6) follows from Theorem 2.2 and the fact that for every Q

$$\left\|\prod_{i=1}^n \left(\frac{p_i\delta_0}{\delta_0 - q_iQ}\right)^{\alpha_i} - \left(\frac{p\delta_0}{\delta_0 - qQ}\right)^{\alpha}\right\| \leqslant \left\|\prod_{i=1}^n \left(\frac{p_i\delta_0}{\delta_0 - q_i\delta_1}\right)^{\alpha_i} - \left(\frac{p\delta_0}{\delta_0 - q\delta_1}\right)^{\alpha}\right\|.$$

Similarly, the bound in (3.7) follows from Theorem 2.4 and for every Q

$$\left\|\prod_{i=1}^n \left(\frac{p_i\delta_0}{\delta_0 - q_iQ}\right)^{\alpha_i} - \exp\big(\lambda(Q-\delta_0)\big)\right\| \leqslant \left\|\prod_{i=1}^n \left(\frac{p_i\delta_0}{\delta_0 - q_i\delta_1}\right)^{\alpha_i} - \exp\big(\lambda(\delta_1-\delta_0)\big)\right\|.$$

Then the proof is completed. ∎

REMARK 3.2. A comparison of the bounds given in (3.6) and (3.7) shows that CNB approximations may be preferred over CP approximations.

3.3. CNB approximation to S_n by Kerstan's method. In this section, we consider the sum S_n of n independent rv's $X_1, X_2, \dots, X_n$ taking values in $\mathbb{R}$. Also, let $p_i = P(X_i \neq 0)$, $q_i = P(X_i = 0)$ and $Q_i(\cdot) = P(X_i \in \cdot \mid X_i \neq 0)$ denote the conditional probability measures. Then, for any Borel measurable set $A \subset \mathbb{R}$,

$$P(X_i \in A) = q_i P(X_i \in A \mid X_i = 0) + p_i Q_i(A) = (q_i \delta_0 + p_i Q_i)(A), \tag{3.8}$$

and hence the distribution of S_n is

$$\mathcal{L}(S_n) = \mathcal{L}(X_1) * \mathcal{L}(X_2) * \ldots * \mathcal{L}(X_n) = \prod_{j=1}^{n} \big(\delta_0 + p_j(Q_j - \delta_0)\big). \tag{3.9}$$

REMARK 3.3. Note that for the representation in (3.9), it suffices that the X_i are independent of S_{i-1} for every $1 \leqslant i \leqslant n$. However, the independence of X_i and S_{i-1}, $1 \leqslant i \leqslant n$, does not imply independence of the X_i (see Example 2.1 of Vellaisamy and Upadhye [27]). We refer to such models as previous-sum independent models. Recently, Vellaisamy and Sankar [26] have used such models for modeling dependent production processes.

Our aim in this section is to approximate the distribution $\mathcal{L}(S_n)$ to a suitable $CNB(n, p, Q) = \mathcal{L}(T_N)$. We choose the parameter p and the distribution Q such that $\mathbb{E}(S_n) = \mathbb{E}(T_N)$ which may possibly reduce $d_{TV}(S_n, T_N)$. Let now

$$Q = \frac{1}{\sum_{i=1}^{n} p_i} \sum_{i=1}^{n} p_i Q_i \tag{3.10}$$

be the probability distribution of Y_1 so that Q is a finite mixture of $Q_1, Q_2, \dots, Q_n$, where Q_j's are as given in (3.8). Since

$$\mathbb{E}(X_i) = \int_{\mathbb{R}} x d(q_i \delta_0 + p_i Q_i) = \int_{\mathbb{R}} x d(p_i Q_i),$$

we have

$$\mathbb{E}(Y_1) = \frac{1}{\sum_{i=1}^{n} p_i} \int_{\mathbb{R}} x d\Big(\sum_{i=1}^{n} p_i Q_i\Big) = \frac{\mathbb{E}(S_n)}{\sum_{i=1}^{n} p_i}.$$

This leads to

$$\mathbb{E}(T_N) = \frac{\mathbb{E}(N)\mathbb{E}(S_n)}{\sum_{i=1}^{n} p_i}.$$

Hence,

$$\mathbb{E}(S_n) = \mathbb{E}(T_N) \Longleftrightarrow \mathbb{E}(N) = \sum_{i=1}^{n} p_i \Longleftrightarrow \frac{nq}{p} = \sum_{i=1}^{n} p_i.$$

Solving for p, we get

$$p = \frac{n}{n + \sum_{i=1}^{n} p_i}. \tag{3.11}$$

Henceforth, all products represent the convolutions. Substituting (3.11) and (3.10) in (3.5), and using (3.9), we obtain

$$\begin{aligned}
(3.12)\quad & \mathcal{L}(S_n) - CNB(n,p,Q) \\
&= \prod_{j=1}^{n} \big(\delta_0 + p_j(Q_j - \delta_0)\big) - \left(\frac{p\delta_0}{\delta_0 - qQ}\right)^n \\
&= \prod_{j=1}^{n} \big(\delta_0 + p_j(Q_j - \delta_0)\big) - \left(\frac{n\delta_0}{n\delta_0 - \sum_{i=1}^{n} p_i(Q_i - \delta_0)}\right)^n \\
&= \big(\prod_{j=1}^{n} (\delta_0 + L_j) - \delta_0\big) CNB(n,p,Q) \\
&= \sum_{j=1}^{n} \sum_{1 \leqslant i_1 < \ldots < i_j \leqslant n} \prod_{s=1}^{j} \big(L_{i_s} CNB(n/j,p,Q)\big),
\end{aligned}$$

where

$$\begin{aligned}
(3.13)\quad L_i &= \big(\delta_0 + p_i(Q_i - \delta_0)\big)\left(\delta_0 - \sum_{j=1}^{n} \frac{p_j}{n}(Q_j - \delta_0)\right) - \delta_0 \\
&= p_i(Q_i - \delta_0) - \big(\delta_0 + p_i(Q_i - \delta_0)\big) \sum_{j=1}^{n} \frac{p_j}{n}(Q_j - \delta_0).
\end{aligned}$$

Therefore,

$$\|\mathcal{L}(S_n) - CNB(n,p,Q)\| \leqslant \sum_{j=1}^{n} \frac{1}{j!} \big(\sum_{i=1}^{n} \|L_i CNB(n/j,p,Q)\|\big)^j, \tag{3.14}$$

where L_i is as defined in (3.13).

3.4. Norm estimates.

LEMMA 3.1. *Let Q and p be defined in* (3.10) *and* (3.11)*, respectively. Then for any $r > 0$*

(i) $$\|(Q_i - \delta_0) CNB(r,p,Q)\| \leqslant 2\left(1 - \frac{p_i}{n+\lambda}\right)^{1\wedge r},$$

(ii) $$\|L_i CNB(r,p,Q)\| \leqslant 2\left(p_i\left(1 - \frac{p_i}{n+\lambda}\right)^{1\wedge r} + \sum_{j=1}^{n} \frac{p_j}{n}\left(1 - \frac{p_j}{n+\lambda}\right)^{1\wedge r}\right),$$

where $\lambda = \sum_{i=1}^{n} p_i$ and $q = 1 - p$.

P r o o f. (i) Let $R_1 = (p_i/\lambda)Q_i$, and $R_2 = \sum_{j=1, j\neq i}^{n} p_j Q_j/\lambda$, so that $Q = R_1 + R_2$. Note first that

$$(Q_i - \delta_0)CNB(r, p, Q) = (Q_i - \delta_0)\left(\frac{p\delta_0}{\delta_0 - qR_1}\frac{\delta_0 - qR_1}{\delta_0 - q(R_1 + R_2)}\right)^r$$

which leads to

$$\|(Q_i - \delta_0)CNB(r, p, Q)\| \leqslant \left\|(Q_i - \delta_0)\left(\frac{p\delta_0}{\delta_0 - qR_1}\right)^r\right\| \|R\|, \tag{3.15}$$

where $R = \big((\delta_0 - qR_1)/(\delta_0 - qQ)\big)^r$. Now

$$\begin{aligned}
\|R\| &= \left\|\left(\frac{\delta_0 - qR_1}{\delta_0 - q(R_1 + R_2)}\right)^r\right\| \\
&= \left\|\left(\frac{\delta_0}{\delta_0 - qR_2/(\delta_0 - qR_1)}\right)^r\right\| \\
&= \left\|\sum_{m=0}^{\infty}\binom{r+m-1}{m}\left(\frac{qR_2}{\delta_0 - qR_1}\right)^m\right\| \\
&\leqslant \sum_{m=0}^{\infty}\binom{r+m-1}{m} q^m \|R_2\|^m \sum_{s=0}^{\infty}\binom{m+s-1}{s} q^s \|R_1\|^s \\
&\leqslant \sum_{m=0}^{\infty}\binom{r+m-1}{m} q^m \left(\frac{\lambda - p_i}{\lambda}\right)^m \sum_{s=0}^{\infty}\binom{m+s-1}{s} q^s \left(\frac{p_i}{\lambda}\right)^s \\
&= \sum_{m=0}^{\infty}\binom{r+m-1}{m} q^m \left(\frac{\lambda - p_i}{\lambda}\right)^m \left(\frac{\lambda}{\lambda - qp_i}\right)^m \\
&= \left(\frac{\lambda - qp_i}{\lambda p}\right)^r.
\end{aligned}$$

Substituting the values of p and q, we get

$$\|R\| \leqslant \left(1 + \frac{\lambda - p_i}{n}\right)^r. \tag{3.16}$$

Consider next

$$\begin{aligned}
&\left\|(Q_i - \delta_0)\left(\frac{p\delta_0}{\delta_0 - qR_1}\right)^r\right\| \\
&\qquad = \left\|(Q_i - \delta_0)\sum_{k=0}^{\infty}\binom{r+k-1}{k} p^r (qR_1)^k\right\| \\
&\qquad = \left\|(Q_i - \delta_0)\sum_{k=0}^{\infty}\binom{r+k-1}{k} p^r \left(\frac{qp_i}{\lambda}\right)^k Q_i^k\right\| \\
&\qquad = p^r\left\|\sum_{k=0}^{\infty}\binom{r+k-1}{k}\mu^k Q_i^{k+1} - \sum_{k=0}^{\infty}\binom{r+k-1}{k}\mu^k Q_i^k\right\|,
\end{aligned}$$

where $\mu = qp_i/\lambda$. Now, putting $k+1=m$ in the first summation and $k=m$ in the second summation, we obtain

$$\begin{aligned}(3.17)\quad &\left\|(Q_i-\delta_0)\left(\frac{p\delta_0}{\delta_0-qR_1}\right)^r\right\| \\ &= p^r\left\|\sum_{m=1}^{\infty}\left(\binom{r+m-2}{m-1}\mu^{m-1}-\binom{r+m-1}{m}\mu^m\right)Q_i^m-\delta_0\right\| \\ &\leqslant p^r\left(1+\sum_{m=1}^{\infty}\left|\binom{r+m-2}{m-1}\mu^{m-1}-\binom{r+m-1}{m}\mu^m\right|\|Q_i\|^m\right) \\ &= p^r\left(1+\sum_{m=1}^{\infty}\binom{r+m-2}{m-1}\mu^{m-1}\left|\frac{m(1-\mu)-(r-1)\mu}{m}\right|\right).\end{aligned}$$

When $r \geqslant 1$, we get from (3.17)

$$\begin{aligned}(3.18)\quad &\left\|(Q_i-\delta_0)\left(\frac{p\delta_0}{\delta_0-qR_1}\right)^r\right\| \\ &\leqslant p^r\left(1+\sum_{m=1}^{\infty}\binom{r+m-2}{m-1}\mu^{m-1}\left((1-\mu)+(r-1)\frac{\mu}{m}\right)\right) \\ &= \frac{2p^r}{(1-\mu)^{r-1}}.\end{aligned}$$

Similarly, when $0<r<1$, we obtain

$$(3.19)\qquad \left\|(Q_i-\delta_0)\left(\frac{p\delta_0}{\delta_0-qR_1}\right)^r\right\| \leqslant 2p^r.$$

Substituting the values of p and μ in (3.18) and (3.19), we finally get

$$\|(Q_i-\delta_0)CNB(r,p,Q)\| \leqslant 2\left(1-\frac{p_i}{n+\lambda}\right)^{1\wedge r},$$

which proves part (i).

(ii) Observe that, from (3.13),

$$\begin{aligned}\|L_iCNB(r,p,Q)\| &\leqslant p_i\|(Q_i-\delta_0)CNB(r,p,Q)\| \\ &\quad+\left\|\big(\delta_0+p_i(Q_i-\delta_0)\big)\sum_{j=1}^{n}\frac{p_j}{n}(Q_j-\delta_0)CNB(r,p,Q)\right\| \\ &\leqslant 2\left(p_i\left(1-\frac{p_i}{n+\lambda}\right)^{1\wedge r}+\sum_{k=1}^{n}\frac{p_k}{n}\left(1-\frac{p_k}{n+\lambda}\right)^{1\wedge r}\right),\end{aligned}$$

using (3.8) and part (i). Hence, the lemma follows. ∎

3.5. The first-order result. We obtain the error bounds on the total variation distance $d_{TV}\big(S_n, CNB(n,p,Q)\big)$ for the choices of p and Q discussed in Subsection 3.4.

THEOREM 3.3. *Let* X_i, $1 \leqslant i \leqslant n$, *be independent real-valued rv's with* $p_i = P(X_i \neq 0)$, *and* $Q_i(\cdot) = P(X_i \in \cdot \mid X_i \neq 0)$. *Also, let*

$$\lambda = \sum_{i=1}^{n} p_i, \qquad Q = \frac{1}{\lambda}\sum_{i=1}^{n} p_i Q_i \quad \text{and} \quad p = \frac{n}{n+\lambda}.$$

Then

$$d_{TV}\big(S_n, CNB(n,p,Q)\big) \leqslant \min\left\{\frac{1}{2}\left(\sum_{j=1}^{n}\frac{\beta_n^j}{j!}\right), 1\right\} \tag{3.20}$$

$$\leqslant \min\{0.911\beta_n, 1\}, \tag{3.21}$$

where $\lambda_2 = \sum_{i=1}^{n} p_i^2$, *and* $\beta_n = 4\big(\lambda - \lambda_2/(n+\lambda)\big)$.

P r o o f. Using part (ii) of Lemma 3.1, we obtain

$$\sum_{i=1}^{n} \|L_i CNB(n/j, p, Q)\| \tag{3.22}$$

$$\leqslant 2\left(\sum_{i=1}^{n} p_i\left(1 - \frac{p_i}{n+\lambda}\right) + \sum_{k=1}^{n} p_k\left(1 - \frac{p_k}{n+\lambda}\right)\right) = 4\left(\lambda - \frac{\lambda_2}{n+\lambda}\right) = \beta_n.$$

Using (3.14) and (3.22), we get

$$\|\mathcal{L}(S_n) - CNB(n,p,Q)\| \leqslant \sum_{j=1}^{n}\frac{1}{j!}\Big(\sum_{i=1}^{n}\|L_i CNB(n/j,p,Q)\|\Big)^j \leqslant \sum_{j=1}^{n}\frac{\beta_n^j}{j!}, \tag{3.23}$$

and hence (3.20) follows.

From part (i) of Lemma 3.1 we infer that

$$d_{TV}\big(S_n, CNB(n,p,Q)\big) \leqslant \min\{f(\beta_n), 1\},$$

where

$$f(x) = \frac{1}{2}\sum_{j=1}^{\infty}\frac{x^j}{j!} = \frac{e^x - 1}{2}.$$

Let $x_0 = \ln(3)$. Then $f(x)$ is increasing, $f(x_0) = 1$, and $f(x) \leqslant x/x_0$ for $x \in (0, x_0)$. Hence, $\min\{f(x), 1\} \leqslant \min\{x/x_0, 1\}$, and so (3.21) follows. ■

3.6. Approximation by a finite signed measure. We consider here the approximation of the distribution of the sum S_n by a finite signed measure defined by

$$W := \big(\delta_0 - \sum_{i=1}^{n} p_i(Q_i - \delta_0)\big)CNB(n,p,Q), \tag{3.24}$$

which is a variant of $CNB(n,p,Q)$ and has the property that $W(\mathbb{R}) = 1$. The choice of this measure is motivated by the expansion of L_i, defined in (3.13), and to remove the first term in the expansion of $\mathcal{L}(S_n) - CNB(n,p,Q)$.

THEOREM 3.4. *Let the assumptions of Theorem* 3.3 *hold*, *and* W *be as defined in* (3.24). *Then*

$$d_{TV}(S_n, W) \leqslant \min\left\{\frac{1}{2}\left(\frac{\beta_n}{2} + \sum_{j=2}^{n}\frac{\beta_n^j}{j!}\right), 1\right\} \tag{3.25}$$

$$\leqslant \min\{0.775\beta_n, 1\}. \tag{3.26}$$

Proof. Consider first

$$\begin{aligned}
&\|\mathcal{L}(S_n) - W\| \\
&\quad = \big\|\mathcal{L}(S_n) - CNB(n,p,Q) - \sum_{i=1}^{n} p_i(Q_i - \delta_0)CNB(n,p,Q)\big\| \\
&\quad = \big\|\sum_{j=1}^{n}\sum_{1\leqslant i_1<\ldots<i_j\leqslant n}\prod_{s=1}^{j}\big(L_{i_s}CNB(n/j,p,Q)\big) \\
&\qquad - \sum_{j=1}^{n} p_j(Q_j - \delta_0)CNB(n,p,Q)\big\|,
\end{aligned} \tag{3.27}$$

using (3.12). Writing the term corresponding to $j = 1$ separately, we get

$$\begin{aligned}
\|\mathcal{L}(S_n) - W\| &= \big\|\sum_{i=1}^{n}\big(L_i - p_i(Q_i - \delta_0)\big)CNB(n,p,Q) \\
&\qquad + \sum_{j=2}^{n}\sum_{1\leqslant i_1<\ldots<i_j\leqslant n}\prod_{s=1}^{j}\big(L_{i_s}CNB(n/j,p,Q)\big)\big\| \\
&\leqslant \sum_{i=1}^{n}\big\|\big(L_i - p_i(Q_i - \delta_0)\big)CNB(n,p,Q)\big\| \\
&\qquad + \sum_{j=2}^{n}\frac{1}{j!}\big(\sum_{i=1}^{n}\|L_iCNB(n/j,p,Q)\|\big)^j.
\end{aligned} \tag{3.28}$$

Since

$$L_i - p_i(Q_i - \delta_0) = -\big(\delta_0 + p_i(Q_i - \delta_0)\big) \sum_{j=1}^{n} \frac{p_j}{n}(Q_j - \delta_0),$$

using Lemma 3.1 (i), we get

$$\begin{aligned}\big\|\big(L_i - p_i(Q_i - \delta_0)\big)CNB(n,p,Q)\big\| &\leqslant \sum_{j=1}^{n} \frac{p_j}{n} \|(Q_j - \delta_0)CNB(n,p,Q)\| \\ &\leqslant \frac{2}{n}\sum_{j=1}^{n}\left(p_j - \frac{p_j^2}{n+\lambda}\right).\end{aligned}$$

Hence,

$$\Big\|\sum_{i=1}^{n}\big(L_i - p_i(Q_i - \delta_0)\big)CNB(n,p,Q)\Big\| \leqslant 2\left(\lambda + \frac{\lambda_2}{n+\lambda}\right) = \frac{\beta_n}{2}.$$

Therefore,

$$\|\mathcal{L}(S_n) - W\| \leqslant \frac{\beta_n}{2} + \sum_{j=2}^{n} \frac{\beta_n^j}{j!}, \tag{3.29}$$

where β_n is as defined in (3.22). Since $W(\mathbb{R}) = 1$ and $\|\mathcal{L}(S_n) - W\| \leqslant 2$, the result in (3.25) follows.

To prove (3.26), note that

$$d_{TV}(S_n, W) \leqslant \min\{g(\beta_n), 1\}, \qquad \text{where } g(x) = \frac{x}{4} + \frac{1}{2}\sum_{j=2}^{\infty}\frac{x^j}{j!}.$$

Note also that $\min\{g(\beta_n), 1\} \leqslant \beta_n/x_0$, where $x_0 \in (0, \infty)$ is the unique solution of $g(x) = 1$. Numerically, it can be seen that $1.29 < x_0 < 1.3$. Therefore,

$$d_{TV}\big(\mathcal{L}(S_n), W\big) \leqslant \min\{0.775\beta_n, 1\}.$$

This proves the theorem. ■

REMARK 3.4. Comparing the practical estimates (3.21) and (3.26), we note that the approximation by a finite signed measure improves the constant of approximation.

4. SOME SPECIAL CASES

In this section, let S_n be defined as in (3.9) and we assume that for every i, $1 \leqslant i \leqslant n$, there exists a probability distribution $\{q_{i,j}\}$ on $\mathbb{N}$ (i.e., $\sum_{j=1}^{\infty} q_{i,j} = 1$)

so that Q_i is a mixture of $\{U_j\}$, a sequence of probability measures concentrated on $\mathbb{R}\backslash\{0\}$. That is,

$$Q_i = \sum_{j=1}^{\infty} q_{i,j} U_j. \tag{4.1}$$

For instance, $q_{i,j} = \delta_{i,j}$, the Kronecker delta, and $U_j = Q_j$ corresponds to the trivial case. Another example due to Roos [20] is the following:

Let $\{B_j\}_{j \geqslant 1}$be a partition of $\mathbb{R}\backslash\{0\}$. Assume $P(X_i \in \cdot \mid X_i \in B_j) = U_j$ is the same for all X_i, $1 \leqslant i \leqslant n$. Then

$$Q_i(\cdot) = \sum_{j=1}^{\infty} P(X_i \in B_j \mid X_i \neq 0) P(X_i \in \cdot \mid X_i \in B_j) = \sum_{j=1}^{\infty} q_{i,j} U_j, \tag{4.2}$$

where $q_{i,j} = P(X_i \in B_j \mid X_i \neq 0)$.

We now require the following lemma.

LEMMA 4.1. *Let* $Q = \sum_{i=1}^{n} p_i Q_i / \lambda$, *where* Q_i *is of the form in* (4.1), *and* $\lambda = \sum_{i=1}^{n} p_i$. *Then for any* $r > 0$ *we have*

$$\|(U_l - \delta_0) CNB(r, p, Q)\| \leqslant 2(1 - q q_l)^{1 \wedge r}, \tag{4.3}$$

$$\|L_i CNB(r, p, Q)\| \leqslant 2 p_i \sum_{l=1}^{\infty} q_{i,l} (1 - q q_l)^{1 \wedge r} + 2 \sum_{j=1}^{n} \frac{p_j}{n} \sum_{m=1}^{\infty} q_{j,m} (1 - q q_m)^{1 \wedge r}, \tag{4.4}$$

where $q_l = \sum_{i=1}^{n} p_i q_{i,l} / \lambda$ *and* $q = \lambda / (n + \lambda)$.

Proof. The proof of the lemma follows along the lines similar to those of Lemma 3.1, except that we now choose $R_1 = q_l U_l$ and $R_2 = \sum_{t=1, t \neq l}^{\infty} q_t U_t$. ■

THEOREM 4.1. *Assume the conditions of Lemma* 4.1 *hold. Let*

$$q_l = \sum_{i=1}^{n} p_i q_{i,l} / \lambda \quad \textit{and} \quad q = \frac{\lambda}{n + \lambda}.$$

Then

$$d_{TV}\big(S_n, CNB(n, p, Q)\big) \leqslant \min\left\{\frac{1}{2} \sum_{j=1}^{n} \frac{\zeta^j}{j!}, 1\right\} \tag{4.5}$$

$$\leqslant \min\{0.911\zeta, 1\}, \tag{4.6}$$

where $\zeta = 4\lambda \left(1 - q \sum_{l=1}^{\infty} q_l^2\right)$.

Proof. The result essentially follows from Lemma 4.1 and the arguments given in the proof of Theorem 3.3. Note that from (4.3) and (4.4), we have

$$\begin{aligned}\sum_{i=1}^{n}\|L_i CNB(n/j,p,Q)\| &\leqslant 4\sum_{l=1}^{\infty}\sum_{i=1}^{n}p_i q_{i,l}(1-qq_l)\\ &= 4\lambda\left(1-\frac{\lambda}{n+\lambda}\sum_{l=1}^{\infty}q_l^2\right)=\zeta.\end{aligned}$$

The practical estimate in (4.6) also follows in a similar manner. ■

Next, we present an analogous result to Theorem 3.4 for the case under consideration, and the proof is omitted.

THEOREM 4.2. *Let W be the signed measure as defined in* (3.24)*. Also, let Q and Q_i be defined in* (3.10) *and* (4.1)*, respectively. Then*

$$d_{TV}(S_n,W) \leqslant \min\left\{\frac{1}{2}\left(\frac{\zeta}{2}+\sum_{j=2}^{n}\frac{\zeta^j}{j!}\right),1\right\} \tag{4.7}$$

$$\leqslant \min\{0.775\zeta,1\}, \tag{4.8}$$

where $\zeta = 4\lambda\left(1-q\sum_{l=1}^{\infty}q_l^2\right)$.

Next, as applications of the above results, we discuss two examples where $Q_i=\sum_{j=1}^{\infty}q_{i,j}U_j$ exists in discrete and continuous cases. Also, we analyze the conditions under which the bounds are optimal.

EXAMPLE 4.1 (Discrete case). Let $\mathcal{L}(Y_i)=Q_i\sim Ge(\eta_i)$, $1\leqslant i\leqslant n$, the geometric distribution with probability distribution (a number of trials for the first success)

$$P(Y_i=k)=(1-\eta_i)^{k-1}\eta_i,\qquad k=1,2,\ldots$$

Let $S_n=\sum_{j=1}^{n}X_j$, where $\mathcal{L}(X_i)=\delta_0+p_i\big(\mathcal{L}(Y_i)-\delta_0\big)$ and $p_i=P(X_i\neq 0)$. Our aim is to approximate $S_n=\sum_{j=1}^{n}X_j$ to $CNB(n,p,Q)$, where p and Q are as defined in (3.11) and (3.10). Let now

$$\eta>\eta_{\max}=\max_{1\leqslant i\leqslant n}\eta_i \quad\text{and}\quad q_{i,j}=(1-b_i)^{j-1}b_i \quad \text{for } j\geqslant 1,$$

where $b_i=\eta_i/\eta$. Choose $U_j=NB(j,\eta)$ with probability distribution

$$U_j(x)=\binom{x-1}{j-1}\eta^j(1-\eta)^{x-j}\quad \text{for } x=j,j+1,\ldots$$

Then it can be easily seen that $Q_i = \sum_{j=1}^{\infty} q_{i,j} U_j$. Using now (4.6), we get

$$\begin{aligned}
&d_{TV}\big(S_n, CNB(n,p,Q)\big)\\
&\quad\leqslant \min\big\{3.65\lambda\big(1 - q\sum_{l=1}^{\infty} q_l^2\big), 1\big\}\\
&\quad= \min\left\{3.65\lambda\left(1 - \frac{\lambda}{n+\lambda}\sum_{l=1}^{\infty}\left(\frac{1}{\lambda}\sum_{i=1}^{n} p_i(1-\eta_i/\eta)^{l-1}\eta_i/\eta\right)^2\right), 1\right\}\\
&\quad\leqslant \min\left\{3.65\lambda\left(1 - \frac{1}{(n+\lambda)\lambda}\sum_{l=1}^{\infty}\sum_{i=1}^{n}\frac{p_i^2(\eta-\eta_i)^{2(l-1)}\eta_i^2}{\eta^{2l}}\right), 1\right\}\\
&\quad= \min\left\{3.65\lambda\left(1 - \frac{1}{(n+\lambda)\lambda}\sum_{i=1}^{n} p_i^2\left(\frac{\eta_i}{2\eta-\eta_i}\right)\right), 1\right\}.
\end{aligned}$$

Note that the above bound is decreasing in η, and so attains the minimum when $\eta = \eta_{\max}$.

EXAMPLE 4.2 (Continuous case). Let $Q_i \sim E(t_i)$, the exponential distribution with density

$$f_{Q_i}(x) = \begin{cases} t_i e^{-t_i x} & \text{for } x > 0,\\ 0 & \text{otherwise,}\end{cases}$$

and $t > \max_{1\leqslant i\leqslant n} t_i$. Let $q_{i,j} = (1-b_i)^{j-1} b_i$, where $b_i = t_i/t$, and $U_j \sim G(t,j)$, the gamma distribution with density

$$f_{U_j}(x|t,j) = \begin{cases} \dfrac{t^j}{(j-1)!} e^{-tx} x^{j-1} & \text{for } x > 0,\\ 0 & \text{otherwise.}\end{cases}$$

Then, it follows that $Q_i = \sum_{j=1}^{\infty} q_{i,j} U_j$. Consequently, from (4.6) we get

$$\begin{aligned}
&d_{TV}\big(S_n, CNB(n,p,Q)\big)\\
&\qquad\leqslant \min\left\{3.65\lambda\left(1 - \frac{\lambda}{n+\lambda}\sum_{l=1}^{\infty} q_l^2\right), 1\right\}\\
&\qquad\leqslant \min\left\{3.65\lambda\left(1 - \frac{1}{(n+\lambda)\lambda}\sum_{i=1}^{n} p_i^2\left(\frac{t_i}{2t-t_i}\right)\right), 1\right\},
\end{aligned}$$

following the arguments in Example 4.1.

Finally, we point out an application of our results to the individual risk model, which is widely used in life and health insurance. Consider a portfolio with n policies with associated non-negative risks, say, $X_1, \ldots, X_n$. Assume that the risk i produces a claim with probability p_i, and let Q_i denote its conditional claim amount. Then $S_n = \sum_{j=1}^n X_j$ denotes the total claim in the individual model. In general, the distribution of S_n is complicated. When all p_i's are small, one may approximate $\mathcal{L}(S_n)$ to a suitable compound distribution (see Roos [20]). If some of the p_i's are not small, it is natural to approximate $\mathcal{L}(S_n)$ to

$$CNB(r, p, Q) = \sum_{k=0}^{\infty} \pi_k(r, p) Q^k, \tag{4.9}$$

where

$$\pi_k(r, p) = \binom{r+k-1}{k} p^r q^k \quad \text{for } k = 0, 1, 2, \ldots,$$

and

$$Q = \frac{1}{\lambda} \sum_{i=1}^{n} p_i Q_i = \sum_{l=1}^{\infty} q_l U_l.$$

Observe that (4.9) is indeed a random sum, and represents the total claim amount in the collective risk model (Grandell [13] or Mikosch [17]). Our results in Theorems 3.3 and 4.1 are helpful to obtain the error estimates in such cases.

Acknowledgments. The authors are grateful to Professors V. Čekanavičius and B. Roos for several comments and encouragements, and also to the referee for his critical comments and suggestions which led to improvements.

REFERENCES

[1] C. D. Aliprantis and O. Burkinshaw, *Principles of Real Analysis*, third edition, Academic Press, San Diego, CA, 1998.
[2] A. D. Barbour, *Multivariate Poisson-binomial approximation using Stein's method*, in: *Stein's Method and Applications*, Singapore Univ. Press, Singapore 2004, pp. 131–142.
[3] A. D. Barbour and P. Hall, *On the rate of Poisson convergence*, Math. Proc. Cambridge Philos. Soc. 95 (3) (1984), pp. 473–480.
[4] A. D. Barbour, L. Holst and S. Janson, *Poisson Approximation*, Oxford Univ. Press, New York 1992.
[5] T. C. Brown and M. J. Phillips, *Negative binomial approximation with Stein's method*, Methodol. Comput. Appl. Probab. 1 (4) (1999), pp. 407–421.
[6] T. C. Brown and A. Xia, *Stein's method and birth-death processes*, Ann. Probab. 29 (3) (2001), pp. 1373–1403.
[7] V. Čekanavičius, *Estimates in total variation for convolutions of compound distributions*, J. London Math. Soc. (2) 58 (3) (1998), pp. 748–760.
[8] V. Čekanavičius and B. Roos, *Two-parametric compound binomial approximations*, Liet. Mat. Rink. 44 (4) (2004), pp. 443–466; translation in: Lithuanian Math. J. 44 (4) (2004), pp. 354–373.
[9] V. Čekanavičius and B. Roos, *Compound binomial approximations*, Ann. Inst. Statist. Math. 58 (1) (2006), pp. 187–210.
[10] V. Čekanavičius and B. Roos, *An expansion in the exponent for compound binomial approximations*, Liet. Mat. Rink. 46 (1) (2006), pp. 67–110; translation in: Lithuanian Math. J. 46 (1) (2006), pp. 54–91.

[11] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, Springer, New York 1988.
[12] S. Drekic and G. E. Willmot, *On the moments of the time of ruin with applications to phase-type claims*, N. Am. Actuar. J. 9 (2) (2005), pp. 17–30.
[13] J. Grandell, *Aspects of Risk Theory*, Springer, New York 1991.
[14] J. Kerstan, *Verallgemeinerung eines Satzes von Prochorow und Le Cam*, Z. Wahrsch. Verw. Gebiete 2 (1964), pp. 173–179.
[15] A. Y. Khintchine, *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, Springer, Berlin 1933.
[16] L. Le Cam, *An approximation theorem for the Poisson binomial distribution*, Pacific J. Math. 10 (1960), pp. 1181–1197.
[17] T. Mikosch, *Non-life Insurance Mathematics*, Springer, Berlin 2004.
[18] H. H. Panjer and G. E. Willmot, *Finite sum evaluation of the negative binomial-exponential model*, Astin Bull. 12 (2) (1981), pp. 133–137.
[19] B. Roos, *On the rate of multivariate Poisson convergence*, J. Multivariate Anal. 69 (1) (1999), pp. 120–134.
[20] B. Roos, *Kerstan's method for compound Poisson approximation*, Ann. Probab. 31 (4) (2003), pp. 1754–1771.
[21] B. Roos, *Poisson approximation of multivariate Poisson mixtures*, J. Appl. Probab. 40 (2) (2003), pp. 376–390.
[22] B. Roos, *Improvements in the Poisson approximation of mixed Poisson distributions*, J. Statist. Plann. Inference 113 (2) (2003), pp. 467–483.
[23] W. Rudin, *Real and Complex Analysis*, third edition, McGraw-Hill, New York 1987.
[24] P. Vellaisamy and B. Chaudhuri, *Poisson and compound Poisson approximations for random sums of random variables*, J. Appl. Probab. 33 (1) (1996), pp. 127–137.
[25] P. Vellaisamy and B. Chaudhuri, *On compound Poisson approximation for sums of random variables*, Statist. Probab. Lett. 41 (2) (1999), pp. 179–189.
[26] P. Vellaisamy and S. Sankar, *A unified approach for modeling and designing attribute sampling plans for monitoring dependent production processes*, Methodol. Comput. Appl. Probab. 7 (3) (2005), pp. 307–323.
[27] P. Vellaisamy and N. S. Upadhye, *On the negative binomial distribution and its generalizations*, Statist. Probab. Lett. 77 (2) (2007), pp. 173–180.
[28] Y. H. Wang, *Coupling methods in approximations*, Canad. J. Statist. 14 (1) (1986), pp. 69–74.
[29] Y. H. Wang, *From Poisson to compound Poisson approximations*, Math. Sci. 14 (1) (1989), pp. 38–49.
[30] H.-J. Witte, *A unification of some approaches to Poisson approximation*, J. Appl. Probab. 27 (3) (1990), pp. 611–621.

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

Received on 1.2.2008;
revised version on 11.11.2008