A Discrete Distribution Arising as a Solution of a Linear Difference Equation:
Extension of the Non Central Negative Binomial Distribution

S. H. Ong, K. Shimizu

* Institute of Mathematical Sciences, University of Malaya, Kuala Lumpur, Malaysia
b Department of Mathematics, Keio University, Yokohama, Japan

Online Publication Date: 01 January 2009
A Discrete Distribution Arising as a Solution of a Linear Difference Equation: Extension of the Non Central Negative Binomial Distribution

S. H. ONG1 AND K. SHIMIZU2

1Institute of Mathematical Sciences, University of Malaya, Kuala Lumpur, Malaysia
2Department of Mathematics, Keio University, Yokohama, Japan

This article considers a discrete distribution that arises as the dominant solution of a linear difference equation. Basic properties and various chance mechanisms that lead to this distribution are given. In particular, its formulation as a weighted distribution and a mixed Poisson process are proposed. Parameter estimation by (a) using a combination of observed frequencies and moments and (b) maximum likelihood are examined. An example of goodness of fit is considered.

Keywords Goodness of fit; Index of dispersion; Mixed, compound, generalized distributions; Mixed Poisson and birth–death with immigration processes; Parameter estimation; Weighted distribution.

Mathematics Subject Classification Primary 60E05; Secondary 62F10, 62P99.

1. Introduction

The derivation of systems of discrete distributions from difference equations have been studied by Katz (1948, 1965) and Ord (1967, 1972), among others. Gurland and Tripathi (1975) and Tripathi and Gurland (1977) have extended the Katz family by relaxing the parameter restrictions in the difference equations. Recently, Kitano et al. (2005) derived a generalized Charlier series distribution from a two-step recursion formula (second-order linear difference equation).

Many mixed Poisson distributions have probability mass functions (pmf) satisfying a second-order linear difference equation. Ong and Muthaloo (1995) examined a mixed Poisson distribution with pmf

\[ P(k) = \frac{(\alpha)_{k}(\beta)_{k}}{k!\gamma^{k}}\psi(k + \alpha, \alpha - \beta + 1; 1/\gamma), \quad k = 0, 1, 2, \ldots, \quad (1.1) \]

Received October 17, 2007; Accepted July 7, 2008
Address correspondence to S. H. Ong, Institute of Mathematical Sciences, University of Malaya, Kuala Lumpur 50603, Malaysia; E-mail: ongsh@um.edu.my
for fitting very long-tailed data, where \( \alpha, \beta, \gamma > 0, \) \( \psi \) denotes the confluent hypergeometric function of the second kind, and \( (x)_k \) Pochhammer’s symbol defined by \( (x)_k = x(x+1) \cdots (x+k-1) \) if \( k \geq 1, (x)_0 = 1. \) The pmf (1.1) satisfies a second-order linear difference equation of the form

\[
(n + 1)(n + 2)y(n + 2) = (n + 1)[2(n + a + 1) + k - c]y(n + 1) - (n + a)(n + a + 1 - c)y(n). \tag{1.2}
\]

This difference equation has two solutions: the dominant and minimal solutions. The pmf (1.1) is a minimal solution (Wimp, 1984) of (1.2). A desirable feature of the minimal solution is that the recurrence relation (1.2) is stable backwards for computing this minimal solution, that is, (1.2) is used to compute \( y(n) \) starting from \( n = N \) where \( N \) is a large integer (for a more detailed account see Ong, 1995; Ong and Muthaloo, 1995). A recurrence relation is said to be stable if the round off errors do not grow relative to the size of the required function.

In this article, we consider the distribution that arises as the dominant solution of the difference equation (1.2). It is found that this distribution is an extension of the non central negative binomial distribution (NNBD) (Ong, 1987; Ong and Lee, 1979, 1986). The NNBD arises as a model in photon and neural counting (see references in Ong and Lee, 1979). It has probability generating function (pgf)

\[
G(z) = (q/(1 - pz))^\nu \exp[\lambda(q/(1 - pz) - 1)], \tag{1.3}
\]

where \( 0 < p = 1 - q < 1, \nu \) and \( \lambda > 0. \) The pmf is given by

\[
P(k) = P(k; p, \nu, \lambda) = e^{-\lambda p} p^k q^\nu L_k^{(\nu-1)}(-\lambda q), \tag{1.4}
\]

where \( L_k^{(\alpha)}(x) \) are the Laguerre polynomials orthogonal over \((0, \infty)\) with respect to \( x^{\alpha-1}e^{-x}. \) The NNBD belongs to the Babel caste \( m \) family of distributions (Letac, 1992). Letac (1992) generalized the NNBD by extending \( \nu > 0 \) to \( \nu \geq 1 - \lambda q. \)

In Sec. 2, the discrete distribution is defined and some basic properties are given. Although the definition of the distribution is mathematically motivated, it is found that a number of chance mechanisms do lead to the distribution making it a potentially useful model for applications. These are discussed in Sec. 3. In particular, we consider its formulation as a weighted NNBD. The idea of weighted distributions, introduced by Rao (1965), has found many practical applications. In Sec. 4, parameter estimation by (a) using a combination of observed frequencies and moments and (b) maximum likelihood are examined. An application of the distribution to data-fitting is considered.

2. A Discrete Distribution: Extension of the NNBD

2.1. Definition

The dominant solution of the difference equation (1.2) (Wimp, 1984) is given by

\[
y(n) = (a)_n {}^1 F_1(n + a, c; x)/n! \quad n = 0, 1, 2, \ldots \tag{2.1}
\]
Here, $\mathbf{1}_{F_1}$ is the confluent hypergeometric function defined by $\mathbf{1}_{F_1}(x, \beta; z) = \sum_{n=0}^{\infty} \frac{(x)_n}{(\beta)_n n!}$. Based upon (2.1), we define the following pmf

$$P(k) = ((v)_k/k!)p^kq^{(x)_k}F_1(v + k, \phi; zq)/F_1(v, \phi; z), \quad k = 0, 1, 2, \ldots$$

(2.2)

where $0 < p = 1 - q < 1, v, \phi$ and $z > 0$. The pgf of (2.2) is

$$G(z) = (q/(1-pz))F_1(v, \phi; zq/(1-pz))/F_1(v, \phi; z), \quad -1 < pz < 1,$$

(2.3)

which is a consequence of using (Erdélyi et al., 1953, p. 283)

$$0^{-a} \sum_{n=0}^{\infty} ((a)_n/n!)(1-0^{-1})^nF_1(a + n, \phi; x) = F_1(a, \phi; 0x), \quad \text{Re} \theta > 1/2.$$

It is easy to check that $P(k) > 0$ and $\sum_{k=0}^{\infty} P(k) = 1$ (since $G(1) = 1$).

The pgf (2.3) shows that the distribution is a convolution of a negative binomial and a stopped-sum hypergeometric probability distribution where

$$G(z) = F_1(v, \phi; z)/F_1(v, \phi; z)$$

(2.4)

is the pgf of a hypergeometric probability distribution (Kemp, 1968) or confluent hypergeometric distribution (Johnson et al., 1992, p. 195; 2005, p. 202). If $v = \phi$, then (2.3) reduces to the NNB pgf (1.3) and if $z$ tends to zero, then (2.3) goes to a negative binomial pgf. The distribution with pgf (2.3) will be referred to as the extended NNBD.

2.2. Recurrence Formulae and Moments

From (1.2), the extended NNBD has probability recurrence relation

$$(k + 1)(k + 2)P(k + 2) = (k + 1)p(2(k + v + 1) + zq - \phi)P(k + 1) - p^2(k + v)(k + v + 1 - \phi)P(k), \quad k \geq 2,$$

(2.5)

where $P(0) = q^vF_1(v, \phi; zq)/F_1(v, \phi; z)$ and $P(1) = vpg^vF_1(v + 1, \phi; zq)/F_1(v, \phi; z)$. It is known that (2.5) is stable forward (Wimp, 1984) for computing $P(k)$.

By differentiating the pgf $r$ times with respect to $z$ and letting $z = 1$, the $r$th descending factorial moment is found to be

$$\mu_r = (v)_r/(p/q)^rF_1(v + r, \phi; \lambda)/F_1(v, \phi; \lambda),$$

(2.6)

and it satisfies the recurrence formula

$$\mu_{r+2} = (p/q)(2(r + v + 1) + \lambda - \phi)\mu_{r+1} - (p/q)^2(r + v)(r + v + 1 - \phi)\mu_r.$$  

(2.7)

The mean and variance are, respectively, given by

$$\mu = \mu_{(1)} = v(p/q)F_1(v + 1, \phi; \lambda)/F_1(v, \phi; \lambda)$$

(2.8)
and
\[ \sigma^2 = v(v + 1)(p/q)^2 i_{1}F_{1}(v + 2, \phi; \lambda)/i_{1}F_{1}(v, \phi; \lambda) + \mu - \mu^2. \tag{2.9} \]

2.3. Series Expansion of the Probabilities

The pmf \( P(k) \) can be expanded in terms of the NNB pmf, giving a relationship between these two distributions.

**Result.** The extended NNB pmf has the following expansion:

\[
P(k) = \frac{(v)}{k!} p^k q^r e^{-2q} \sum_{n=0}^{\infty} \frac{(v + k)}{(\phi)_n} P(n; q, \phi, \lambda)/i_{1}F_{1}(v, \phi; \lambda),
\]

where \( P(n; q, \phi, \lambda) \) is the NNB pmf given by (1.4) with \( p \) and \( v \) replaced by \( q \) and \( \phi \), respectively. \( \square \)

**Proof.** The expansion is a result of the formula (Buchholz, 1969, p. 138, Eq. (11))

\[
{i_{1}F_{1}(a, c; xy(x - 1))} = (1 - x)^a \sum_{n=0}^{\infty} \frac{(a)_n}{(c)_n} L_n^{(c-1)}(y)x^n, \quad |x| < 1
\]

on setting \( a = v + k, c = \phi, x = q, \) and \( y = -\lambda p. \) \( \square \)

Similarly, the constant \( {i_{1}F_{1}(v, \phi; \lambda)} \) has the expansion

\[
{i_{1}F_{1}(v, \phi; \lambda)} = \left( \frac{1}{2} \right)^{-c} e^q \sum_{n=0}^{\infty} \frac{(v)_n}{(\phi)_n} P(n; 1, \phi, 2\lambda).
\]

2.4. Index of Dispersion (ID)

The value of \( \text{ID}(X) = \text{Var}(X)/E(X) \) is obtainable from the mean and the variance given by (2.8) and (2.9), but it is not clear if \( \text{ID} \) is less than 1 (underdispersion) or greater than 1 (overdispersion). We use the formulation as a mixed NB distribution (Sec. 3.3.1) to show that the distribution with pmf (2.2) is overdispersed.

If \( K \) has a confluent hypergeometric distribution with pgf (2.4), then from

\[
\frac{d^n}{dz^n} i_{1}F_{1}(a, b; z) = \frac{(a)_n}{(b)_n} i_{1}F_{1}(a + n, b + n; z)
\]

it follows that

\[
E(K) = (v\phi/\lambda)i_{1}F_{1}(v + 1, \phi + 1; \lambda)/i_{1}F_{1}(v, \phi; \lambda),
\]

\[
E(K(K - 1)) = [z^2v(v + 1)/{\phi(\phi + 1)}]i_{1}F_{1}(v + 2, \phi + 2; \lambda)/i_{1}F_{1}(v, \phi; \lambda).
\]

If \( Z \) has a geometric distribution with pgf \( q/(1 - pt) \), then the mean and the variance of \( Z \) are given by \( E(Z) = p/q, \text{Var}(Z) = p/q^2 \), and we have \( \text{Var}(Z) > E(Z) \). The mean and the variance of the compound variable \( S = Z_1 + Z_2 + \cdots + Z_K \) can be obtained from

\[
E(S) = E(Z)E(K), \quad \text{Var}(S) = E(K)\text{Var}(Z) + \text{Var}(K)\{E(Z)\}^2
\]
and it follows from $\text{Var}(Z) > E(Z)$ that $\text{Var}(S) > E(K)E(Z) = E(S)$. Thus, we have

$$\text{ID}(S) = \text{Var}(S)/E(S) > 1.$$  

In Sec. 3.3.1, the rv with pmf (2.2) may be represented by the sum $X = X_1 + S$, where $X_1$ has the negative binomial with pgf $(q/(1 - pt))^r$ and $S$ is the above compound variable. The fact that the distribution of $X$ is over dispersed follows from $\text{Var}(X_1) = vp/q^2 > E(X_1) = vp/q$ and $\text{Var}(S) > E(S)$.

**Remark 2.1.** In general, the fact that the variance is larger than the mean may be seen as follows: Since $X = X_1 + S$ is NB for a given $S = s$, we have

$$\text{var}(X + s) = (k + s)q/p^2 > E(X) = (k + s)(q/p)$$

and if $s$ varies as a positive random variable $S$, the inequality always holds.

### 3. Models for the Extended NNBD

#### 3.1. A Shifted Weighted NNBD

Let $X$ be a random variable with pmf $p(k)$, and assume that the probability of ascertaining the event $X = k$ has a weighting factor $w(k)$. The pmf of the resulting distribution is then

$$P(k) = P(X = k) = w(k)p(k)/\sum w(k)p(k). \quad (3.1)$$

This is known as the weighted distribution (Rao, 1965, 1985) with weight $w(k)$.

In what follows, we shall derive the extended NNBD as a weighted NNBD. Consider the NNB pmf $p(k) = e^{-\lambda p}p^k q^r k^{-1}(-\lambda q)$ and weight $w(k) = k^r$, one of the weights deemed to be useful in scientific work (Patil et al., 1986). Then the weighted distribution has pmf

$$P(k) = k^r p(k)/\left\{ r!(p/q)^r L_{r-1}^r(-\lambda) \right\}, \quad k \geq r,$$

because

$$\sum w(k)p(k) = r!(p/q)^r L_{r-1}^r(-\lambda)$$

is the $r$th descending factorial moment of the NNBD (Ong and Lee, 1979). Let $k = x + r$, that is, shift $k$ to 0. We have

$$P(x) = \frac{(r + 1)x}{x!} p^x q^r e^{-\lambda p} \frac{L_{r-1}^r(-\lambda q)}{L_{r-1}^r(-\lambda)}, \quad x \geq 0. \quad (3.2)$$

Pmf (3.2) is found to be pmf (2.2) if the confluent hyper geometric definition of the Laguerre polynomial

$$L_1^x(y) = \left[(x+1)/k!\right]_1 F_1(-k, x+1; y)$$
is used together with Kummer’s transformation

\[ \text{$_1F_1(x, \beta; y) = e^y \text{$_1F_1(\beta - x, \beta; -y)$}} \]  

(3.3)

and making appropriate substitution.

**Remark 3.1.** In the pmf (3.2), \( r \) need not be an integer. If \( r \) is replaced by a positive real number \( \gamma \), then (3.2) becomes

\[ P(x) = \frac{(\gamma + 1)^x}{x!} p^x q^{x+y} e^{-\lambda p} L^{\gamma-1}_{\gamma+1}(\lambda q) L^{\gamma-1}_{\gamma+1}(\lambda^2). \]

When \( \gamma = 0 \), we get the NNB pmf. Thus the parameter \( \gamma \) measures the deviation of the distribution from the NNBD.

### 3.2. Conditional Model

Consider the Type II bivariate NNBD (Lee and Ong, 1986) with joint pgf

\[ G(u, v) = \left( \frac{q}{1-p_1 u - p_2 v} \right)^v \exp \left\{ \lambda \left( \frac{q}{1-p_1 u - p_2 v} - 1 \right) \right\}, \]

where \( 0 < q = 1 - p_1 - p_2 < 1 \), \( v, \lambda > 0 \). The conditional pgf is given by

\[ G_{Y|X}(z) = \left( \frac{1-p_2}{1-p_2 z} \right)^z \text{$_1F_1(x + z, \lambda q/(1-p_2 z))$} / \text{$_1F_1(x + z, \lambda q/(1-p_2))$}, \]

where \( x = v + x \) and \( \Lambda = \lambda q/(1-p_2) \). This is of the form (2.3).

### 3.3. Mixture Models

The mixture models given in this section follow those of the NNBD (Ong, 1987; Ong and Lee, 1979) and so the details will be left out.

#### 3.3.1. Mixed Negative Binomial and Mixed NNBD.

**Mixed negative binomial.** Let \( X \mid k \) be a negative binomial random variable with parameters \((p, v + k)\) such that \( k \) is a hypergeometric probability distribution random variable (Johnson et al., 1992, p. 84; 2005, p. 88) with pgf (2.4). The unconditional pgf is then given by (2.3).

**Mixed NNBD.** Let \( X \mid k \) be a NNB random variable with parameters \((p, \phi + \gamma + k, \lambda)\) such that \( k \) has a hypergeometric probability distribution with pgf

\[ G_k(z) = \text{$_1F_1(-\gamma, \phi; -\lambda q/(1-p z))$} / \text{$_1F_1(-\gamma, \phi; -\lambda)$}. \]
The unconditional pgf of $X$ is given by

$$G_X(z) = G_1(z) G_2(z) = (q/(1 - pz))^{\phi+\gamma} F_1(\phi + \gamma, \phi; \lambda q/(1 - pz))/F_1(\phi + \gamma, \phi; \lambda),$$

where

$$G_1(z) = (q/(1 - pz))^\phi \exp\{\lambda q/(1 - pz) - 1\} \quad \text{and} \quad G_2(z) = (q/(1 - pz))^{\gamma} F_1(-\gamma, \phi; -\lambda q/(1 - pz))/F_1(-\gamma, \phi; -\lambda)$$

after applying Kummer’s transformation (3.3). If we let $v = \phi + \gamma$, then $G_X(z)$ is the pgf (2.3). When $\gamma$ goes to 0, $G_X(z)$ is the NNB pgf.

3.3.2. Compound Distribution. If $Z_1, Z_2, \ldots$ are mutually independent and identically distributed Bernoulli rv’s with each pgf $p_1 t + q_1 (0 < p_1 < 1, p_1 + q_1 = 1)$ and $K$ is independent of $Z_1, Z_2, \ldots$ and is distributed with pgf (2.3), then the pgf of the compound variable $S = Z_1 + Z_2 + \cdots + Z_K$ has the same type as (2.3):

$$G(z) = (Q/(1 - Pz))^{\gamma} F_1(v, \phi; \lambda Q/(1 - Pz))/F_1(v, \phi; \lambda)$$

where

$$P = pp_1/(1 - pq_1), \quad Q = q/(1 - p q_1).$$

It is difficult to find an explicit form for the pmf of $S$ if the distribution of $Z$’s is different from a Bernoulli distribution. In that case one can use the recursive algorithm of Kitano et al. (2005).

3.4. Mixed Poisson Distribution

Let $X | \theta$ be a Poisson random variable conditional on the parameter $\theta$ such that $\theta$ is a gamma-type random variable with probability density function

$$f(\theta) = (a^\theta \theta^{r-1} \exp(-a \theta)/\Gamma(r))_0 F_1(\theta; \phi; a \lambda \theta)/_1 F_1(v, \phi; \lambda),$$

which is a special case of (see Johnson et al., 1992, p. 339; 2005, p. 379)

$$g(y) = \frac{e^{-y} y^{c-1} p F_q(a_1, a_2, \ldots, a_p; b_1, \ldots, b_q; \theta y)}{\Gamma(c) p+1 F_q(a_1, a_2, \ldots, a_p, c; b_1, \ldots, b_q; \theta)}.$$  \hspace{1cm} (3.5)

The unconditional pmf of $X$

$$P(k) = \int_0^\infty e^{-\theta k} \frac{\theta^k}{k!} f(\theta) d\theta$$  \hspace{1cm} (3.6)

is given by (2.2).

Note. In this case, the mixing of a generalized hypergeometric probability distribution (Poisson) with a gamma-type distribution (3.5) does not produce another generalized hypergeometric probability distribution as stated in Johnson et al. (1992, p. 339; 2005, p. 379).
3.5. Mixed Poisson Process and Birth–Death with Immigration Process

The application of the mixed Poisson distribution usually occurs as a mixed Poisson process, that is, with equation (3.6) given as

\[ P(k) = \int_0^\infty e^{-\theta t} \frac{(\theta t)^k}{k!} f(\theta) d\theta, \]

where \( k \) is the number of events in a time interval \((0, t)\). For this case, \( p = (t + \lambda)^{-1} \) in the pmf (2.2). Corresponding to this mixed Poisson process it may be shown that its intensity function (see, for instance, McFadden, 1965) is given by

\[ r_n(t) = \frac{p(v + n)}{t} \frac{\Gamma(v + n + 1, \lambda q)}{\Gamma(v + n, \lambda q)}. \]

Let

\[ \gamma_n(t) = \frac{t}{p(n + 1)} r_n(t). \]

Then, due to the recurrence relation (1.2), we have

\[ n(n + 1)\gamma_{n+1}(t) = n [2(n + v + \lambda q - \phi)] \gamma_n(t) - (n + v - 1)(n + v - \phi)\gamma_{n-1}(t), \]

and this facilitates the computation of \( r_n(t) \). The mixed Poisson process may be characterized as a birth process (McFadden, 1965, p. 89) where \( r_n(t)dt \) represents the conditional probability of a birth in \((t, t + dt)\) given \( n \) previous births in \((0, t)\).

The extended NNBD also arises as a birth–death with immigration process:

Consider the birth–death with immigration process having birth rate \( \lambda \), death rate \( \mu \), immigration rate \( \nu \), and distribution of initial population size \( \gamma_i \), \( i = 0, 1, 2, \ldots \). If \( \gamma_i \) has the hypergeometric probability distribution with pgf (2.4), then the birth–death with immigration process has the extended NNBD (for details, see Ong, 1987).

4. Parameter Estimation and Goodness of Fit

4.1. Estimation Based Upon Observed Frequencies and First Two Moments

Due to the complicated expressions for the probabilities and moments, estimation by the usual method of moments and zero frequency will give estimates of the parameters requiring iterative solution. We shall consider the simpler three-parameter distribution with \( v = \phi + \gamma \) where \( \gamma \) is known. For this three-parameter case, it is possible to obtain simple estimates of the parameters. The recurrence relations (2.5) and (2.7) for the probabilities and factorial moments respectively give the following equations:

\[ 2P(2) = p(2(\phi + \gamma + 1 + \lambda q - \phi)P(1) - p^2(\phi + \gamma)(\gamma + 1)P(0), \]

\[ 6P(3) = 2p(2(\phi + \gamma + 2 + \lambda q - \phi)P(2) - p^2(\phi + \gamma + 1)(\gamma + 2)P(1), \]

\[ \mu(2) = (p/q)(2(\phi + \gamma + 1 + \lambda - \phi)\mu(1) - (p/q)^2(\phi + \gamma)(\gamma + 1). \]
Straightforward but lengthy calculations lead to the equations

\[ p^2(\phi + \gamma) = \left[ 6f_3f_1 - 4f_2^2 - 4pf_1f_2 + p^2(\gamma + 2)f_1^2 \right] / \left[ 2(\gamma + 1)f_0f_2 - (\gamma + 2)f_1^2 \right], \quad (4.4) \]

\[ ap^2 + bp + c = 0, \quad (4.5) \]

where

\[ a = -f_1\hat{\mu}_{(2)} + f_1\hat{\mu}_{(1)}((\gamma + 2)f_1(-f_1 + (\gamma + 1)f_0)/\delta - (\gamma + 2)) - (\gamma + 1)(\gamma + 2)f_1^2/\delta, \]

\[ b = 2\left[ f_1\hat{\mu}_{(2)} + 2(f_1/\delta) \left\{ \hat{\mu}_{(1)}f_2f_1 - (\gamma + 1)(f_0\hat{\mu}_{(1)} - f_1)f_2 \right\} \right], \]

\[ c = 2f_2\hat{\mu}_{(1)} - f_1\hat{\mu}_{(2)} - \left( \frac{6f_3f_1 - 4f_2^2}{\delta} \right) \left\{ f_1\hat{\mu}_{(1)} - (\gamma + 1)(f_0\hat{\mu}_{(1)} - f_1) \right\}, \]

\[ \delta = 2(\gamma + 1)f_0f_2 - (\gamma + 2)f_1^2, \]

\( f_i, i = 0, 1, 2, 3 \) are the observed proportions of zeroes, ones, twos, and threes, and \( \hat{\mu}_{(i)}, i = 1, 2 \) are the sample descending factorial moments. The solutions of (4.5) give rise to two estimates of \( p \). In this case, the usual procedure is to select the estimate \( \tilde{p} \) which satisfies \( 0 < \tilde{p} < 1 \) or which gives a better chi-square goodness of fit. For this estimate \( \tilde{p} \), (4.4) gives

\[ \tilde{\phi} = \left[ 6f_3f_1 - 4f_2^2 - 4\tilde{p}f_1f_2 + \tilde{p}^2(\gamma + 2)f_1^2 \right] / \left[ 2(\gamma + 1)f_0f_2 - (\gamma + 2)f_1^2 \right] \tilde{p}^2 - \gamma \]

while (4.3) results in

\[ \tilde{\lambda} = \left[ \hat{\mu}_{(2)} - (\tilde{p}/\tilde{q})(2(\tilde{\phi} + \gamma + 1) - \tilde{\phi})\hat{\mu}_{(1)} + (\tilde{p}/\tilde{q})^2(\tilde{\phi} + \gamma)(\gamma + 1) \right] / \left[ (\tilde{p}/\tilde{q})\hat{\mu}_{(1)} \right]. \]

### 4.2. Maximum Likelihood Estimation

Although simple estimates \( (\tilde{p}, \tilde{\phi}, \tilde{\lambda}) \) may be obtained for the parameters with \( \gamma \) known, these estimates may be out of the range or inefficient. An efficient method of estimation is given by maximum likelihood estimation (MLE). Let \( a_k \) be the observed frequency of \( k \) events and \( t \) the highest value of \( k \) observed. The sample mean is

\[ \bar{x} = \sum_{k=0}^{t} ka_k/N, \]

where \( N = \sum_{k=0}^{t} a_k \) is the sample size. Let the likelihood function \( L(p, \phi, \lambda) \) be given by

\[ L = \prod_{k=0}^{t} (P(k))^{a_k}. \]

The solutions to

\[ \frac{\partial}{\partial \theta} \ln L = \sum_{k=0}^{t} a_k \frac{1}{P(k)} \frac{\partial}{\partial \theta} P(k) = 0, \]
where \( \theta \) denotes one of the parameters \( p, \phi, \) or \( \lambda \), give the ML estimates. The partial derivatives of \( P(k) \) with respect to the parameters are summarized below.

**Result.** Let \( \varphi(\lambda) = \left[ \psi_1(\phi + \gamma, \phi; \lambda) \right]^{-1} \) and \( \varphi'(\lambda) = d\varphi(\lambda)/d\lambda \). Then

\[
\frac{\partial}{\partial p} P(k) = \frac{(-k + 1)P(k + 1) + kP(k)}{pq}, \tag{4.6}
\]

\[
\frac{\partial}{\partial \lambda} P(k) = \frac{\varphi'(\lambda)}{\varphi(\lambda)} P(k) + \frac{1}{\lambda} \left\{ (k + 1)P(k + 1)/p - (k + \phi + \gamma)P(k) \right\}, \tag{4.7}
\]

\[
\frac{\partial}{\partial \phi} P(k) = P(k) \ln q + \sum_{i=0}^{k} P(i)\frac{p^{k-i}}{(k-i)} + \varphi(\lambda) \sum_{i=0}^{\infty} \frac{(\phi + \gamma)_i}{(\phi)_i} (S(i; k) - P(k)) D(i) \frac{\lambda^i i!}{i!}, \tag{4.8}
\]

where

\[
S(i; k) = \binom{\phi + \gamma + i + k - 1}{k-1} p^k q^{\phi+\gamma+i} \quad \text{and} \quad D(i) = \sum_{j=1}^{i} \left( \frac{1}{\phi + \gamma + j - 1} - \frac{1}{\phi + j - 1} \right).
\]

**Proof.** The partial derivatives (4.6)–(4.8) are easily determined from the partial derivatives of the pgf \( G(z) \) by extracting the \( k \)th term. In particular, for the partial derivative (4.8), the following results have been used:

\[
\frac{\partial}{\partial a} (a)_n = (a)_n \left\{ \psi(a + n) - \psi(a) \right\},
\]

\[
\frac{\partial}{\partial \phi} F_1(\phi + \gamma, \phi; y) = \sum_{i=0}^{\infty} \frac{(\phi + \gamma)}{(\phi)_i} \left[ \psi(\phi + \gamma + 1) - \psi(\phi + 1) \right] \left[ \psi(\phi + i) - \psi(\phi + 1) \right] y^i i!,
\]

where \( \psi(z) \) is the logarithmic derivative of the gamma function or digamma function and

\[
\psi(z + n) = \psi(z) + \sum_{j=1}^{n} (z + j - 1)^{-1}, \quad n = 1, 2, 3, \ldots,
\]

(see, for instance, Johnson et al., 1992, p. 7; 2005, p. 9). From these partial derivatives of \( P(k) \) the log-likelihood equations give

\[
\sum_{k=0}^{t} a_k (k + 1)P(k + 1) / P(k) = N\bar{x}, \quad \bar{x} = -(p/q)[\phi + \gamma - \lambda^{-1}\varphi(\lambda)/\varphi(\lambda)],
\]

\[
N \ln q + \sum_{k=0}^{t} a_k \sum_{i=0}^{k} P(i) \frac{p^{k-i}}{(k-i)} + \varphi(\lambda) \sum_{k=0}^{t} a_k \sum_{i=0}^{\infty} \frac{(\phi + \gamma)_i}{(\phi)_i} (S(i; k) - P(k)) D(i) \frac{\lambda^i i!}{i!} = 0. \tag{4.9}
\]

The system of equations in (4.9) may be solved by the Newton–Raphson method to give the ML estimates \((\hat{p}, \hat{\lambda}, \hat{\phi})\) with \( \gamma \) known. If \( \gamma \) is unknown, it may be varied to determine the maximum of the likelihood function (method of profile likelihood).
Table 1

Thunderstorm events at Cape Kennedy, Florida June (Falls et al., 1971)

<table>
<thead>
<tr>
<th>No.</th>
<th>Observed frequency</th>
<th>NBD</th>
<th>NNBD</th>
<th>Extended NNBD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>187</td>
<td>181.68</td>
<td>186.21</td>
<td>186.99</td>
</tr>
<tr>
<td>1</td>
<td>77</td>
<td>87.54</td>
<td>80.95</td>
<td>77.09</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>36.80</td>
<td>37.04</td>
<td>39.86</td>
</tr>
<tr>
<td>3</td>
<td>17</td>
<td>14.71</td>
<td>15.68</td>
<td>16.86</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>5.73</td>
<td>6.29</td>
<td>6.22</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>2.20</td>
<td>2.42</td>
<td>2.08</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1.34</td>
<td>1.41</td>
<td>0.90</td>
</tr>
<tr>
<td>Total</td>
<td>330</td>
<td>330.00</td>
<td>330.00</td>
<td>330.00</td>
</tr>
</tbody>
</table>

\[ x^2 = 2.15 \quad 0.75 \quad 0.024 \]

Degrees of freedom

<table>
<thead>
<tr>
<th></th>
<th>NBD</th>
<th>NNBD</th>
<th>Extended NNBD</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNBD: ( \hat{\rho} = 0.2403, \hat{v} = 0.0001, \hat{\lambda} = 2.3811 ).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extended NNBD: ( \hat{\rho} = 0.1023, \hat{v} = 0.0235, \hat{\phi} = 5.1605, \hat{\lambda} = 14.7001 ).</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.3. *An Example of Goodness of Fit*

To illustrate the application of the extended NNBD, we consider the fit of the distribution to the data on thunderstorm events at Cape Kennedy, Florida for the month of June (Falls et al., 1971) in Table 1. The ML estimates \( (\hat{\rho}, \hat{v}, \hat{\phi}, \hat{\lambda}) \) have been determined by numerical optimization of the log-likelihood function instead of solving (4.9). The NBD and NNBD have also been fitted for comparison. The expected frequencies have not been pooled to calculate the chi-square values. It is seen that the extended NNBD provides a good fit. For the fit by NNBD, ML estimate \( \hat{v} \) is close to 0 indicating that its special case, the Pólya–Aeppli distribution, is the more appropriate model.

**Acknowledgment**

The authors wish to thank the referee and Associate Editor for their insightful comments which led to a greatly improved article.

**References**


