c(θ)-Geometric Distribution of Order k and Some of its properties

Okoli, O.C.

Department of Mathematics, Chukwuemeka Odumegwu Ojukwu University, Uli Campus Anambra State Nigeria

*Corresponding author’s email: odicomatics@yahoo.com, odico2003@yahoo.com

ABSTRACT
In this research work, we present a new distribution function that modifies and generalizes the classical geometric distribution of order k via c(θ)-geometric distribution which we called the c(θ)-geometric distribution of order k (simply θ-geometric distribution of order k if θ ≥ 0 and reverse θ-geometric distribution of order k if θ < 0.) for every θ ∈ ℤ. The associated cumulative distribution, generating functions and statistics were also established. The results obtained in this paper improve, generalize and complement the works of several authors in the literature.

1. Introduction and Preliminary

Geometric distribution is one of the discrete probability functions that one often comes across in most standard text and journal papers on discrete distributions. This distribution has received attention of several researchers over the years, as an introductory part of this work, it is important we consider the following definition for the standard geometric distributions.

Definition 1.1
A random variables X and Y is said to have a (standard) geometric distribution if the probability mass function is given by

\[ f(x) = (1 - q)q^{x-1}; \quad x = 1, 2, \ldots; \quad 0 \leq q \leq 1 \]  

or

\[ f(y) = (1 - q)q^y; \quad y = 0, 1, \ldots; \quad 0 \leq q \leq 1 \]  

with the corresponding cumulative distribution

\[ F(x) = 1 - q^x \]  

and

\[ F(x) = 1 - q^{y+1} \]  

Observe that (1.1) yields equation (1.2) if we take \( y = x - 1 \), hence equation (1.2) is a transformation of equation (1.1). The random variable \( X \) in equation (1.1) and the random \( Y \) in equation (1.2) are often described as; the number of independent (Bernoulli’s) trials on which the first success occurs and the number of failures before the first success occurs.
2. Geometric Distribution of order k

Philippou et al. (1983) introduced the distribution of the number of trials until the first occurrence of consecutive k successes in Bernoulli trials with success probability p, which they called a geometric distribution of order k. This distribution has caught the interest of researchers, owing to the work of Philippou et al. (1983) whose contribution in this field seems to be very important. Philippou et al. (1983) studied a geometric distribution of order k and defined a Poisson distribution of order k and a negative binomial distribution of order k, attention has been paid to interrelationships among the so-called discrete distributions of order k and their exact distribution theory has been developed extensively by many researchers.

Ever since then, related concepts have been introduced, reliability theory of the consecutive-k-out-of-n: F systems and studied by Kontoleon (1980), Chiang and Niu (1981) and Derman, Lieberman and Ross (1982). Several related concepts on discrete distributions of order k were studied by Lambiris and Papastavridis (1985), Fu (1985;1986a;1986b), Fu and Beihua (1987), Aki (1985), Hirano (1986), Philippou (1986), Papastavridis (1987), Aki and Hirano (1988; 1989;1996;1997), Chrysaphinou and Papastavridis (1990c), Fu and Koutras (1994a;1994b), Godbole (1990a;1991), Griffith (1983), Hirano and Aki (1993), Charalambides (1994) and Makri and Philippou (1994), Feller (1968). Various modifications have been made also on the underlying sequence, e.g., replacing the Bernoulli trials for other random sequences such as some urn models, a binary sequence of order k as in Aki (1985), Aki and Hirano (1994;1995), Dhar and Jiang (1995) and Balakrishnan (1997). For any fixed positive integer k, let \( N_k \) denote the number of independent trials with success probability \( p \) (\( q = 1 - p \)) until the first occurrence of consecutive k successes. Philippou and Muwafi (1982) gave the following definition for the probability mass function of the event \( \{ N_k = x \} \)

**Definition 2.1**
A random variable \( N_k \) is said to have a geometric distribution of order k with parameter \( p \) if the probability mass function (pmf) is given by

\[
f_k(x) = P(N_k = x) = \sum_{x_1 + 2x_2 + \ldots + kx_k = x - k} \binom{x}{x_1, x_2, \ldots, x_k} \times p^{x_1} q^{x_2} \cdots q^{x_k}; \quad x \geq k \quad (1.5)
\]

Where the summation is taken over all non-negative integers \( x_1, x_2, \ldots, x_k \) satisfying the condition \( x_1 + 2x_2 + \ldots + kx_k = x - k \).

Observe that equation (1.5) reduces to the standard geometric distribution in equation (1.2) if we take \( k = 1 \), this implies that the generalized geometric distributions of order k which was introduced by Philippou et al. (1983) which has been studied by some of the above mentioned authors does not directly generalized the two standard geometric distributions \( f(x) \) and \( f(y) \) given in equation (1.1) and equation (1.2) simultaneously. Thus the distribution due to Philippou et al. (1983) in equation (1.5) is rather restrictive. In this research work, in addition to removing-filling this gap of restriction, we shall:

(a). Introduce an infinite families \( \Lambda \) of standard geometric distributions that properly contains the probability mass function (pmf) in equation (1.1) and equation (1.2).

(b). Define an associated generalized geometric distribution of order k in relation to the infinite families \( \Lambda \) of standard geometric distributions and its properties.

(c). Investigate its relationship with other discrete distributions.

3. \( c(\theta) \)-Geometric Distribution

In this section, we introduce the \( c(\theta) \)-geometric distribution. Let \( N_\theta \) be a discrete random variable, \( Z \) the set of integers, then for every \( \theta \in Z \) we denote the associated \( c(\theta) \)-geometric distribution by \( f_\theta(x; p) \) where \( c(\theta) \) is an integer function of \( \theta \), and then we define \( \Lambda = \{ f_\theta(x; p) : \theta \in Z \} \) with \( f_\theta(x; p) = P(N_\theta = x) \).

**Definition 3.1**
For any given \( \theta \in Z \), a random variable \( N_\theta \) is said to have a \( c(\theta) \)-geometric distribution with parameter \( p \) if the probability mass function (pmf) is given by

\[
f_\theta(x; p) = pq^{x-c(\theta)}; \quad x \geq c(\theta) \quad (1.6)
\]

Where \( c(\theta) = 1 - \theta \), we shall simply call \( f_\theta(x; p) \), a \( \theta \)-geometric distribution if \( \theta \geq 0 \) and reverse \( \theta \)-geometric distribution if \( \theta < 0 \). The corresponding cdf is given by
\[
F_0(x) = \begin{cases} 
0 & \text{if } x < c(\theta) \\
1 - q^{x-c(\theta)+1} & \text{if } c(\theta) \leq x < \infty \\
1 & \text{if } x \geq \infty
\end{cases}
\]  

(1.7)

Observe that \( \lim F_0(x) = 1 \forall \theta \in \mathbb{Z} \). Consequently \( f_0 \) is a well-defined pmf. And if we take \( \theta = 0 \) and \( \theta = 1 \), we obtain the standard results in equation (1.1) to equation (1.4) as special cases. By implication, the zero \((0)\)-geometric distribution, implies that the first success occurs at the \( x \)th trial and one \((1)\)-geometric distribution implies that the first success occurs at the \((x + 1)th\) trial. Hence, in general, the random variable \( N_\theta \) describes (counts) the number of independent trials on which the first success occurs at least \( \theta \)-steps before the \( x \)th trial or at most \( \theta \)-steps after the \( x \)th trial.

### 4 \((c(\theta))\)-Geometric Distribution of order \( k \).

Motivated by the work of Philippou et al., (1983) in definition 2.1 and the definition of \( c(\theta)\)-geometric distribution we introduced in definition 3.1. Furthermore, we introduce yet another new distribution in the definition that follows. However, in the success-failure \( \{S,F\} \) (binary) sequence with probability \([p,q]\) we require that for every \( \alpha \in \{p,q\}, \theta \neq 0,k = 1 \) with respect to \( c(\theta,k) \), that the following condition \((A)\) holds:

\[
\alpha^{c(\theta,k)} = \begin{cases} 
p, & \alpha = p \\
q^{c(\theta,k)}, & \alpha = q
\end{cases}
\]

The essence of this condition \((A)\) is to ensure that the \( k \) consecutive success is achieve for every \( \theta \neq 0 \).

#### Definition 4.1

For any given \( \theta \in \mathbb{Z} \) and \( k \in \mathbb{N} \), a random variable \( N_{0,k} \) is said to have a \( c(\theta)\)-geometric distribution of order \( k \) with parameter \( p \) denoted by \( G_{0,k}(.;p) \) if the probability mass function \((pmf)\) is given by

\[
f_{0,k}(x) = P(N_{0,k} = x) = \sum_{x=\infty}^{\infty} \binom{x}{x_1, x_2, \ldots, x_k} \left(x_{0,1} + x_{0,2} + \cdots + x_{0,k}\right) \times
\]

\[
p^x \left(\frac{q}{p}\right)^{x_0,1 + x_0,2 + \cdots + x_0,k}; \ x \geq c(\theta,k)
\]

(1.8)

Where \( c(\theta,k) = k - \theta \). If we put \( \theta = 0 \) in equation (1.8) we obtain equation (1.5) as a special case.

#### Theorem 1.3

Let \( m = m(t) \) be a continuous, real-valued function, suppose there exists \( t^* \in \mathbb{R} \) which is a zero of \( m(t) - 1 \), then the generating function \((g.f.)\) of \( N_{0,k} \) distributed as \( G_{0,k}(.;p) \) is given by

\[
G_{N_{0,k}}(m(t)) = \frac{(mp)^{k-\theta}(1-mp)}{1 - mp^k}; |m(t)| \leq 1 \forall t \in \mathbb{R}
\]

(1.9)

**Proof.**

For simplicity, we put \( m = m(t) \)

\[
E(m^x) = \sum_{x=\infty}^{\infty} m^x P(N_{0,k} = x)
\]

\[
= \sum_{x=\infty}^{\infty} \sum_{x_0,1, x_0,2, \ldots, x_0,k} \left(x_{0,1} + x_{0,2} + \cdots + x_{0,k}\right) \times
\]

\[
= \frac{(mp)^{k-\theta}(1-mp)}{1 - mp^{k+1}}
\]

Observe that, one consequence of theorem 1.3 is that, if \( m(t) \) is the identity function; that is \( m(t) = t \) then (1.9) becomes the probability generating function \((pgf)\) and if \( m(t) = e^t \), \( m(t) = e^{\theta t} \), then (1.9) becomes the moment generating function, characteristic function of the distribution respectively.

#### Remark 1.4

As a corollary to theorem 1.3, observe that if we put \( \theta = 0 \) and \( m(t) = t \) we have \( G_{N_{0,k}}(t) = \frac{(pt)^k(1-pt)}{1-\theta qt(pt)} \) which has been obtained by Feller (1968), Philippou et al. (1983; 1984), Aki (1994), Shao and Wang (2015) via different approach. In the remaining part of this work, if we put \( \theta = 0 \) then the resulting corollaries are result obtained by some of the above
mentioned authors. We allow the reader to infer that, due to want of space and conciseness we may not state all these resulting corollaries.

**Theorem 1.5** Let \( N_{0,k} \) be as in theorem 1.3 and \( m(t) = e^t \), Then

i. \[ E(N_{0,k}) = \frac{p^n}{(qp^k)^3} \sum_{j=1}^{\alpha} a_j(qp^k)^j \]

ii. \[ Var(N_{0,k}) = \frac{p^n}{(qp^k)^3} \sum_{j=0}^{3} \beta_j(qp^k)^j - \left( \frac{p^n}{(qp^k)^3} \right)^2 \left( \sum_{j=1}^{\alpha} a_j(qp^k)^j \right)^2 \]

**Proof.**

(i) If we put \( \eta = k - \theta \), observe that \( M_{N_{0,k}}(t) = ((pe^t)^n - (pe^t)^{n+1})(1 - e^t + qp^ke^{tk+1})^{-1} \), so that

\[
\frac{dM_{N_{0,k}}(t)}{dt} = (\eta(p)e^t - (n + 1)(pe^t)^{n+1})(1 - e^t + qp^ke^{tk+1})^{-1} - ((pe^t)^n - (pe^t)^{n+1})(1 - e^t + qp^ke^{tk+1})^{-2}(-e^t + (k + 1)qp^ke^{tk+1})
\]

Hence

\[
\left. \frac{dM_{N_{0,k}}(t)}{dt} \right|_{t=0} = \frac{1}{(qp^k)^2} \left[ \eta qp^k + (n + 1)qp^k + p^2 - (k + 1)q \right] = \frac{p^n}{(qp^k)^3} \sum_{j=1}^{2} a_j(qp^k)^j ; \quad (1.10)
\]

Where

(ii) Observe that

\[
\left. \frac{d^2M_{N_{0,k}}(t)}{dt^2} \right|_{t=0} = \frac{1}{(qp^k)^3} \left[ 2\eta rp^k + (n + 1)p^2 - (k + 1) \right] + \frac{1}{(qp^k)^3} \left[ -2(n + 1)p^2 + (k + 1) \right] = \frac{p^n}{(qp^k)^3} \sum_{j=0}^{3} \beta_j(qp^k)^j ; \quad (1.11)
\]

where

\[
\beta_0 = 2q, \quad \beta_1 = (2\eta + 1 + 4k - 2n), \quad \beta_2 = (2\eta + 1 + 4k - 2n)^2, \quad \beta_3 = (k + 1)^2 + (2\eta + 1) + (k + 1)^2
\]

Using equation (1.10) and equation (1.11), and the fact that

\[
Var(N_{0,k}) = \left. \frac{d^2M_{N_{0,k}}(t)}{dt^2} \right|_{t=0} - \left( \frac{dM_{N_{0,k}}(t)}{dt} \right) \left|_{t=0} \right)^2
\]

The result follows immediately.

**Corollary 1.6** Let \( N_{0,k} \) be as in theorem 1.5. and we put \( \theta = 0 \) then

i. \[ E(N_{0,k}) = \frac{1-p^k}{qp^k} \]

ii. \[ Var(N_{0,k}) = \frac{1-(2(k+1)qp^k+p^k+1)}{q^2p^k} \]

**Proof.**

We shall prove (i), where (ii) follows similarly. Recall we have in equation (1.10) that
\[E(N_{0,k}) = \frac{p^n}{(qp^k)^2} \sum_{j=1}^{2} \alpha_j(qp^k)^j\]

\[= \frac{1}{(qp^k)^2} [qp^{k+1}q + q^2p^{2k+1}(nq - p - (k + 1)q)] = \frac{[1 - (1 + p)q + \theta p^k]}{qp^{k+1}q + q^2p^{2k+1}}\]

Hence substituting \(\theta = 0\) yield the required result.

**Definition 1.7** Let \(N^1_{0,k}, N^2_{0,k}, \ldots, N^r_{0,k}\) be independent identically distributed random variables. Then the distribution of \(N_{0,k} = N^1_{0,k} + N^2_{0,k} + \cdots + N^r_{0,k}\) is called a \(c(0)\)-negative binomial distribution of order \(k\) denoted by \(N_{0,k}(c); p)\)
The following theorem is a consequence of definition 1.6.

**Theorem 1.8** Let \(N^1_{0,k}, N^2_{0,k}, \ldots, N^r_{0,k}\) be independent identically distributed random variables. Then the probability generating function of \(N_{0,k} = N^1_{0,k} + N^2_{0,k} + \cdots + N^r_{0,k}\)

\[G_{N_{0,k}+1/k+1,r}(t) = \left(\frac{(pt)^{k-0}}{1 - \frac{q}{p} \sum_{j=1}^{r}(pt)^j}\right)^r; \ |t| \leq 1, \quad (1.13)\]

**Proof.**
Since \(N^1_{0,k}, N^2_{0,k}, \ldots, N^r_{0,k}\) are independent identically distributed random variables with \(p, g, f. G_{N^i_{0,k}}\) (as in (1.6)) for each \(j \in \{1,2,\ldots,r\}\), thus it follows that

\[G_{N_{0,k}}(t) = \prod_{j=1}^{r} G_{N^j_{0,k}}(t)\]

\[= \left(G_{N^j_{0,k}}(t)\right)^r = \left(\frac{(pt)^{k-0}}{1 - \frac{q}{p} \sum_{j=1}^{r}(pt)^j}\right)^r\]

**Theorem 1.9** Let \(N_{0,k}\) be as in definition 1.7, then the probability \(\{N_{0,k} = x\}\) is given by

\[f_{0,k,r}(x) = P(N_{0,k} = x) = \sum_{x_0,1 \geq 0 \cdots x_0,k \geq 0} x_0,1 + x_0,2 + \cdots + kx_0,k = x - c(0,k)r\]

Where the summation is taken over all non-negative integers \(x_0,1, x_0,2, \ldots, x_0,k\) such that \(x_0,1 + 2x_0,2 + \cdots + kx_0,k = x - c(0,k)r\).

**Proof.**
By theorem 1.8 the \(p, g, f.\) of \(N_{0,k}\) is given by

\[G_{N_{0,k}}(t) = \left(\frac{(pt)^{k-0}}{1 - \frac{q}{p} \sum_{j=1}^{r}(pt)^j}\right)^r\]

Hence by binomial expansion we have that

\[G_{N_{0,k}+1/k+1,r}(t) = ((pt)^{k-0})^r \sum_{x=0}^{\infty} \binom{x}{r} \left(-\frac{q}{p} \sum_{j=1}^{r}(pt)^j\right)^x\]

\[= ((pt)^{k-0})^r \sum_{x=0}^{\infty} \frac{(x + r - 1)!}{(r - 1)!} \sum_{x_0,1+x_0,2+\ldots+x_0,k=x} x_0,1 + x_0,2 + \cdots + x_0,k (pt)^{x_0,1+x_0,2+\ldots+x_0,k} (pt)^{x_0,1+x_0,2+\ldots+x_0,k} (\frac{q}{p})^{x_0,1+x_0,2+\ldots+x_0,k}\]

\[= ((pt)^{k-0})^r \sum_{x=0}^{\infty} \frac{(x + r - 1)!}{(r - 1)!} \sum_{x_0,1+x_0,2+\ldots+x_0,k=x} x_0,1 + x_0,2 + \cdots + x_0,k (pt)^{x_0,1+x_0,2+\ldots+x_0,k} (\frac{q}{p})^{x_0,1+x_0,2+\ldots+x_0,k}\]

\[= ((pt)^{k-0})^r \sum_{x=0}^{\infty} \frac{(x + r - 1)!}{(r - 1)!} \sum_{x_0,1+x_0,2+\ldots+x_0,k=x} x_0,1 + x_0,2 + \cdots + x_0,k (pt)^{x_0,1+x_0,2+\ldots+x_0,k} (\frac{q}{p})^{x_0,1+x_0,2+\ldots+x_0,k}\]

\[= ((pt)^{k-0})^r \sum_{x=0}^{\infty} \frac{(x + r - 1)!}{(r - 1)!} \sum_{x_0,1+x_0,2+\ldots+x_0,k=x} x_0,1 + x_0,2 + \cdots + x_0,k (pt)^{x_0,1+x_0,2+\ldots+x_0,k} (\frac{q}{p})^{x_0,1+x_0,2+\ldots+x_0,k}\]
\[= \sum_{x=(k-0)r}^{\infty} t^x P(N_{0,k,r} = x)\]

Hence the result follows immediately.

**Definition 1.10.** Let \(N_{0,k,\mu}\) be a discrete random variables and \(\mu\) a nonnegative integer constant. Assuming that \(q \to 0\) and \(rq \to \mu\) as \(r \to \infty\). Then \(N_{0,k,r} \to N_{0,k,\mu}\). We say that \(N_{0,k,\mu}\) has the \(c(\theta)\)-Poisson distribution of order \(k\) denoted by \(P_{0,k}(\cdot; \mu)\).

**Theorem 1.11** Let \(N_{0,k,\mu}\) be as in definition 1.10, then the probability \(\{N_{0,k,\mu} = x\}\) is given by

\[
\lim_{r \to \infty} P(N_{0,k,r} - (k - \theta)r = x) = \sum_{x_0,x_1,x_2,\ldots,x_{k-1} \geq 0, x_k \geq x - (k-0)r} e^{-\mu(k-\theta)} \frac{\mu^x}{x!} \prod_{j=1}^{k} x_j! \forall i_m \geq 0 \quad (1.15)
\]

**Proof.**

It suffices to show that the characteristic function of \(N_{0,k,r} - (k - \theta)r\) converges to \(P_{0,k}(\cdot; \mu)\) as \(r \to \infty\). From (1.9) we have

\[
\phi_{N_{0,k,r}}(t) = \left(\frac{1 - q(\mu e^{it})^k}{1 - \frac{q}{p} \sum_{j=1}^{k} (\mu e^{it})^j}\right)^r
\]

Since \(|e^{it}|^{k-\theta} \leq 1\) for \(|t| \leq 1\). It then follows that

\[
\phi_{N_{0,k,r}}(t) \leq \left(\frac{p^{k-\theta}}{1 - \frac{q}{p} \sum_{j=1}^{k} (\mu e^{it})^j}\right)^r = \left(1 - \frac{\mu}{pr} \sum_{j=1}^{k} (\mu e^{it})^j\right)^r
\]

By hypothesis, we have that \(q \to 0\) and \(rq \to \mu\) as \(r \to \infty\). Thus, as \(r \to \infty\) we have that

\[
\lim_{r \to \infty} \phi_{N_{0,k,r}}(t) \leq \lim_{r \to \infty} \left(1 - \frac{\mu}{pr} \sum_{j=1}^{k} (\mu e^{it})^j\right)^r = \exp\left(-\mu(k-\theta) + \sum_{j=1}^{k} \mu e^{itj}\right)
\]

Hence the result follows immediately.

**Conclusion**

we presented a new distribution function that modifies and generalize the classical geometric distribution of order \(k\) via \(c(\theta)\)-geometric distribution which we called the \(c(\theta)\)-geometric distribution of order \(k\) (simply \(\theta\)-geometric distribution of order \(k\) if \(\theta \geq 0\) and reverse \(\theta\)-geometric distribution of order \(k\) if \(\theta < 0\)) for every \(\Theta \in \mathbb{Z}\). The associated cumulative distribution, generating functions and statistics are also established. The results obtained in this paper improve, generalize and complement the works of several authors in the literature.

**References**


**How to cite this article**