ON A GENERALIZATION OF LING’S BINOMIAL DISTRIBUTION

Femin YALCIN
Department of Mathematics, Faculty of Arts and Sciences
Izmir University of Economics, 35330, Balcova, Izmir, Turkey

Abstract: In a sequence of \( n \) binary trials, distribution of the random variable \( M_{n,k} \), denoting the number of overlapping success runs of length exactly \( k \), is called Ling’s binomial distribution or Type II binomial distribution of order \( k \). In this paper, we generalize Ling’s binomial distribution to Ling’s \( q \)-binomial distribution using Bernoulli trials with a geometrically varying success probability. An expression for the probability mass function of this distribution is derived. For \( q = 1 \), this distribution reduces to Ling’s binomial distribution.

Key words: Binomial distribution of order \( k \), Ling’s binomial distribution, \( q \)-distributions, runs.

History: Submitted: 13 July 2013; Revised: 13 September 2013; Accepted: 1 October 2013

1. Introduction

Much attention has been paid to the distribution of the number of runs of fixed length, say \( k \), in a sequence of \( n \) (\( n \geq k \)) binary trials. Ling (1988) [4] introduced a binomial distribution of order \( k \) which is based on overlapping counting. This distribution is also called Type II binomial distribution of order \( k \) and differs from the Type I binomial distribution of order \( k \), which was studied by Hirano (1986) [3] and Philippou and Makri (1986) [6]. The Type I binomial distribution of order \( k \) is based on nonoverlapping counting scheme. In a sequence of \( n \) binary trials, we use \( N_{n,k} \) (\( M_{n,k} \)) to denote a random variable which has Type I (Type II) binomial distribution of order \( k \). \( N_{n,k} \) (\( M_{n,k} \)) is actually the number of nonoverlapping (overlapping) success runs of length exactly \( k \) in \( n \) trials. Consider a sequence of \( n = 12 \) trials 10110111110 assuming “1 ” as a success and “0” as a failure. If \( k = 3 \), then \( N_{12,3} = 2 \) and \( M_{12,3} = 4 \).

Ling (1988) [4] obtained the following recursive and nonrecursive equations for the pmf of \( M_{n,k} \) when the corresponding binary trials are independent and identically distributed with the probability of success \( p \).

\[
P\{M_{n,k} = x\} = \begin{cases} 
p^n & \text{if } x = n - k + 1, \\
2p^{n-1}q & \text{if } x = n - k (>0), \\
\sum_{j=1}^{x+k} p^{j-1}qP\{M_{n-j,k} = x - \max(0, j-k)\} & \text{if } 0 \leq x < n - k 
\end{cases}
\]

and

\[
P\{M_{n,k} = x\} = \sum_{i=0}^{n} \sum_{\text{max}(0,i-k+1)+\sum_{j=k+1}^{n} (j-k)x_j=x}^{x_i} \binom{x_1 + x_2 + \cdots + x_n}{x_1, x_2, \ldots, x_n} p^n \left( \frac{q}{p} \right)^{\sum_{i=1}^{n} x_i}.
\]
Then, Godbole (1992) \cite{godbole1992} derived a simpler formula

\[
P \{ M_{n,k} = x \} = \begin{cases} 
  p^n & \text{if } x = n - k + 1, \\
  \sum_{y=\left\lceil \frac{n}{2} \right\rceil}^{n} q^y p^{n-y} \sum_{j=0}^{\left\lceil \frac{n-y}{2} \right\rceil} (-1)^j \binom{n+y}{j} \binom{n-j}{y-k} & \text{if } x = 0, \\
  \sum_{y=\left\lceil \frac{n-k}{2} \right\rceil}^{n-k-1} q^y p^{n-y} \sum_{i=1}^{y-n+k} \binom{y+1}{i} \binom{i-1}{x} & \text{if } 1 \leq x \leq n - k.
\end{cases}
\]

An even simpler formula was obtained by Makri, Philippou, and Psillakis (2007) \cite{makri2007}. For \( s = 0 \) and \( l = k - 1 \), Theorem 2.1. of \cite{makri2007} gives

\[
P \{ M_{n,k} = x \} = \begin{cases} 
  \sum_{y=\left\lceil \frac{n}{2} \right\rceil}^{n} q^y p^{n-y} C(n-y+y+1,k-1) & \text{if } x = 0, \\
  \sum_{y=\left\lceil \frac{n-k}{2} \right\rceil}^{n-k-(k-1)} q^y p^{n-y} \sum_{i=1}^{y-n+k} \binom{y+1}{i} \binom{i-1}{x} & \text{if } x \neq 0,
\end{cases}
\]

where

\[
C(\alpha; i, r - i; m - 1, n - 1) = \sum_{j_1=0}^{\left\lceil \frac{\alpha}{r} \right\rceil} \sum_{j_2=0}^{\left\lfloor \frac{\alpha - m j_1}{r} \right\rfloor} (-1)^{j_1+j_2} \binom{i}{j_1} \binom{r-i}{j_2} \binom{\alpha - m j_1 - n j_2 + r - 1}{r - 1}
\]

and

\[
C(\alpha, r; m - 1) = C(\alpha; i, r - i; m - 1, m - 1) = \sum_{j=0}^{\left\lfloor \frac{\alpha}{r} \right\rfloor} (-1)^j \binom{r}{j} \binom{\alpha - m j + r - 1}{r - 1}.
\]

Charalambides (2010) \cite{charalambides2010} studied discrete \( q \)-distributions on Bernoulli trials with a geometrically varying success probability. Let us consider a sequence \( X_1, \ldots, X_n \) of zero (failure)-one (success) Bernoulli trials such that the trials of the subsequences after the \((i-1)\)st zero until the \( i \)th zero are independent with failure probability

\[
q_i = 1 - \theta q^{i-1}, \quad i = 1, 2, \ldots, 0 < \theta < 1, 0 < q \leq 1.
\]

The probability mass function of the number \( Z_n \) of successes in \( n \) trials \( X_1, \ldots, X_n \) is given by

\[
P \{ Z_n = r \} = \binom{n}{r} \theta^r \prod_{i=1}^{n-r} (1 - \theta q^{i-1})
\]

for \( r = 0, 1, \ldots, n \), where

\[
\binom{n}{r} = \frac{[n]_q!}{[r]_q! [n-r]_q!}
\]

and \( [x]_q = [x]_q [x-1]_q \cdots [x-k+1]_q, \quad [x]_q = (1-q^x)/(1-q), \quad [x]_q! = [1]_q [2]_q \cdots [x]_q \). The distribution given by 1.5 is called a \( q \)-binomial distribution.
Yalcin and Eryilmaz (2014) [7] obtained the distribution of $N_{n,k}$ for the model 1.4. The resulting distribution is the Type I $q$-binomial distribution of order $k$. In this paper, we study the distribution of $M_{n,k}$ under the model 1.4. The new distribution is called Type II $q$-binomial distribution of order $k$ or Ling’s $q$-binomial distribution.

Note that, throughout the paper, for integers $n$ and $m$, and real number $x$, let $\binom{n}{m}$ and $\lfloor x \rfloor$ denote the binomial coefficients and the greatest integer less than or equal to $x$, respectively. We also assume for convenience that if $a > b$, then $\sum_{i=a}^{b} = 0$ and $\prod_{i=a}^{b} = 1$.

2. Type II $q$-binomial distribution of order $k$

We first note the following Lemma which will be useful in the sequel.

**Lemma 1.** For $0 < q \leq 1$, define

$$B_q(r,s,t) = \sum_{y_1 + \cdots + y_r = s} \sum_{y_j \geq 0, j=1,2,\ldots,r} q^{y_2 + 2y_3 + \cdots + (r-1)y_r},$$

where

$$I_j = \begin{cases} 1 & \text{if } y_j \geq k, \\ 0 & \text{otherwise} \end{cases}$$

and $y_j$s are nonnegative integers, $j = 1,2,\ldots,r$. Then $B_q(r,s,t)$ obeys the following recurrence relation

$$B_q(r,s,t) = \begin{cases} \sum_{j=0}^{k-1} q^{(r-1)j} B_q(r-1,s-j,t) & \text{if } r > 1, s \geq 0, \text{ and } t \geq 0, \\ + \sum_{j=k}^{s} q^{(r-1)j} B_q(r-1,s-j,t+j+k-1) & \text{if } (r = 1, s \geq k \text{ and } t = s-k+1) \text{ or } (r = 1, 0 \leq s < k \text{ and } t = 0), \text{ otherwise.} \\ 1 & \text{otherwise.} \end{cases}$$

**Proof.** Considering the values that $y_r$ can take, we have

$$B_q(r,s,t) = \sum_{y_1 + \cdots + y_{r-1} = s-1} \sum_{y_j \geq 0, j=1,2,\ldots,r-1} q^{y_2 + 2y_3 + \cdots + (r-2)y_{r-1}} + q^{r-1} \sum_{y_1 + \cdots + y_{r-1} = s-2} \sum_{y_j \geq 0, j=1,2,\ldots,r-1} q^{y_2 + 2y_3 + \cdots + (r-2)y_{r-1}}$$

$$+ \cdots + q^{(k-1)(r-1)} \sum_{y_1 + \cdots + y_{r-1} = s-k+1} \sum_{y_j \geq 0, j=1,2,\ldots,r-1} q^{y_2 + 2y_3 + \cdots + (r-2)y_{r-1}} + q^{k(r-1)} \sum_{y_1 + \cdots + y_{r-1} = s-k} \sum_{y_j \geq 0, j=1,2,\ldots,r-1} q^{y_2 + 2y_3 + \cdots + (r-2)y_{r-1}}.$$
Thus the proof is completed. □

**Theorem 1.** For 0 < q ≤ 1, the probability mass function of the number of overlapping success runs of length k in n trials is given by

\[
P\{M_{n,k} = x\} = \sum_{i=0}^{n} \theta^{n-i} \prod_{j=1}^{i} (1 - \theta q^{j-1}) B_q(i+1, n-i, x),
\]

(2.1)

\(x = 0, 1, \ldots, n-k+1.\)

**Proof.** Let \(S_n\) denote the total number of zeros (failures) in n binary trials. Then

\[P\{M_{n,k} = x\} = \sum_{i} P\{M_{n,k} = x, S_n = i\}.\]

The joint event \(\{M_{n,k} = x, S_n = i\}\) can be described with the following binary sequence which consists of \(i\) zeros.

\[\underbrace{1\ldots011\ldots0}_{y_1}1\ldots01\ldots1,\]

where

\[y_1 + y_2 + \cdots + y_{i+1} = n - i\]

s.t.

\[I_1(y_1 - k + 1) + I_2(y_2 - k + 1) + \cdots + I_{i+1}(y_{i+1} - k + 1) = x\]

\[y_j \geq 0 \text{ and } I_j = \begin{cases} 1 & \text{if } y_j \geq k, \\ 0 & \text{otherwise}, \end{cases} \quad j = 1, 2, \ldots, i+1.\]

Under the model 1.4,

\[P\{M_{n,k} = x\} = \sum_{i} \sum_{\substack{y_1 + \cdots + y_{i+1} = n-i \\text{ s.t.} \\frac{y_1+\cdots+y_{i+1}+y_j}{y_j} = x}} (\theta q^0)^{y_1} (1 - \theta q^0) (\theta q)^{y_2} (1 - \theta q) \cdots (\theta q^{i-1})^{y_i} \times (1 - \theta q^{i-1}) (\theta q^i)^{y_{i+1}} \]

\[= \sum_{i=0}^{n} \theta^{n-i} \prod_{j=1}^{i} (1 - \theta q^{j-1}) \sum_{\substack{y_1 + \cdots + y_{i+1} = n-i \\text{ s.t.} \\frac{y_1+\cdots+y_{i+1}+y_j}{y_j} = x}} q^{y_2+2y_3+\cdots+(i-1)y_{i+1}} \]

\[= \sum_{i=0}^{n} \theta^{n-i} \prod_{j=1}^{i} (1 - \theta q^{j-1}) B_q(i+1, n-i, x).\]

Thus the proof is completed. □
Remark 1. For $q = 1$, the probability mass function of $M_{n,k}$ given in 2.1 is an alternative to 1.1, 1.2, and 1.3 because the Type II 1-binomial distribution of order $k$ is actually Ling’s binomial distribution.

Example 1. For $n = 5$ and $k = 2$, below we compute the pmf of $M_{n,k}$.

\[
P \{ M_{5,2} = 0 \} = \theta^3 (1 - \theta) (1 - \theta q) q^3 + \theta^2 (1 - \theta) (1 - \theta q) (1 - \theta q^2) (q + q^2 + 2q^4 + q^5) \\
+ (1 - \theta) (1 - \theta q) + (1 - \theta q^2) (1 - \theta q^3) (1 + q + q^2 + q^3 + q^4) \\
+ (1 - \theta) (1 - \theta q) (1 - \theta q^2) (1 - \theta q^3) (1 - \theta q^4), \\
P \{ M_{5,2} = 1 \} = \theta^4 (1 - \theta) (1 - \theta q) (q + q^2 + 2q^4 + q^5) \\
+ (1 - \theta) (1 - \theta q) (1 - \theta q^2) (1 + q + q^2 + q^3 + q^4), \\
P \{ M_{5,2} = 2 \} = \theta^5 (1 - \theta)(1 - \theta q) (q + q^2 + q^3) + \theta^3 (1 - \theta) (1 - \theta q) (1 + q^3 + q^6), \\
P \{ M_{5,2} = 3 \} = \theta^4 (1 - \theta) (1 + q^4), \\
P \{ M_{5,2} = 4 \} = \theta^5.
\]

In Table 1 and Table 2, we respectively compute the probability mass function of $M_{10,2}$ for selected values of the parameters $\theta$ and $q$ and the expected value of $M_{n,k}$ for different choices of $k, n$ and the parameters $\theta$ and $q$. The numerical results indicate that $E(M_{n,k})$ is increasing in both $\theta$ and $q$.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$\theta = 0.5, q = 0.5$</th>
<th>$\theta = 0.5, q = 0.8$</th>
<th>$\theta = 0.9, q = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.68854</td>
<td>0.48212</td>
<td>0.14144</td>
</tr>
<tr>
<td>1</td>
<td>0.16216</td>
<td>0.24141</td>
<td>0.08809</td>
</tr>
<tr>
<td>2</td>
<td>0.07597</td>
<td>0.13410</td>
<td>0.07780</td>
</tr>
<tr>
<td>3</td>
<td>0.03699</td>
<td>0.07168</td>
<td>0.06976</td>
</tr>
<tr>
<td>4</td>
<td>0.01828</td>
<td>0.03682</td>
<td>0.06291</td>
</tr>
<tr>
<td>5</td>
<td>0.00913</td>
<td>0.01880</td>
<td>0.05812</td>
</tr>
<tr>
<td>6</td>
<td>0.00453</td>
<td>0.00833</td>
<td>0.05202</td>
</tr>
<tr>
<td>7</td>
<td>0.00244</td>
<td>0.00465</td>
<td>0.06236</td>
</tr>
<tr>
<td>8</td>
<td>0.00098</td>
<td>0.00111</td>
<td>0.03882</td>
</tr>
<tr>
<td>9</td>
<td>0.00098</td>
<td>0.00098</td>
<td>0.34868</td>
</tr>
</tbody>
</table>

Table 1. Probability mass function of $M_{10,2}$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$k$</th>
<th>$\theta = 0.5, q = 0.5$</th>
<th>$\theta = 0.5, q = 0.8$</th>
<th>$\theta = 0.9, q = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2</td>
<td>0.6048</td>
<td>1.0661</td>
<td>5.1925</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.2721</td>
<td>0.4288</td>
<td>4.2369</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>0.6067</td>
<td>1.1416</td>
<td>7.3807</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.2734</td>
<td>0.4485</td>
<td>6.3441</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.0638</td>
<td>0.0866</td>
<td>4.8372</td>
</tr>
</tbody>
</table>

Table 2. Expected value of $M_{n,k}$.

Acknowledgements

The author would like to express her sincere thanks to the anonymous referee for careful reading of the first version of this paper and his/her useful comments and suggestions.

References


