Statistics and Applications {ISSN 2454-7395 (online)}
Volume 21, No. 2, 2023 (New Series), pp 259–278
http://www.ssca.org.in/journal

Number of Overlapping Runs Until a Stopping Time for Higher Order Markov Chain

Anuradha
Department of Statistics, Lady Shri Ram College for Women
University of Delhi, Lajpat Nagar-IV, New Delhi – 110024, India

Received: 02 November 2022; Revised: 23 December 2022; Accepted: 05 April 2023

Abstract

Let $\{X_j : j \geq -m+1\}$ be a homogeneous Markov chain of order m taking values in $\{\mathbf{0}, \mathbf{1}\}$. For $j = 0, -1, \ldots, -l+1$, we will set $R_j = 0$ and we define $R_j = \prod_{i=j-1}^{j-l}(1 - R_i) \prod_{i=j}^{j+k-1} X_i$. Now $R_j = 1$ implies that an l-look-back run of length k has occurred starting at j. Here R_j is defined inductively as a run of $\mathbf{1}$'s starting at j, provided that no l-look-back run of length k occurs, starting at time $j-1, j-2, \ldots, j-l$ respectively. We study the conditional distribution of the number of overlapping runs of length k_1 until the stopping time *i.e.* the r^{th} occurrence of the l-look-back run of length k where $k_1 \leq k$ and obtain it's probability generating function. The number of overlapping runs of length k_1 until the stopping time has been expressed as the sum of r independent random variables with the first random variable having a slightly different distribution. We introduce a new discrete distribution, namely *generalized Binomial type* distribution, which plays a central role in our study. The conditional distributions are identified using this and other known distributions, such as extended negative binomial distribution of order k. Our results also generalize the known results for the number of successes until a stopping time.

Key words: Overlapping runs; Stopping time; Markov chain; Strong Markov property; Probability generating functions.

AMS Subject Classifications: 60C05, 60E05, 60F05

1. Introduction

Since Feller (1968) introduced runs of successes as an example of a renewal event, the theory of distributions of runs has been explored widely by the researchers. The application of powerful techniques such as Markov embedding technique (see Fu and Koutras (1994)), method of conditional p.g.f.s (see Ebneshahrashoob and Sobel (1990)) *etc.* has paved way to develop and study the distributions of various run statistics and their properties extensively.

Two schemes of counting runs, namely non-overlapping counting and the overlapping counting, have been extensively studied in the literature. As the name suggests, in the

Corresponding Author: Anuradha
Email: sarkar.anuradha@gmail.com

non-overlapping counting, runs are not allowed to overlap while in the overlapping counting scheme the runs may overlap as much as possible. Philippou and Makri (1986) studied the distribution of number of non-overlapping runs of successes of length k for *i.i.d.* Bernoulli trials and introduced the Binomial distribution of order k. Ling (1988) derived the distribution of the number of overlapping runs of successes of length k for a sequence of *i.i.d.* Bernoulli trials. This distribution is referred as the Type II Binomial distribution of order k. Aki and Hirano (1994) obtained the marginal distributions of number of failures, successes and success-runs of length less than k until the first occurrence of consecutive k successes when the underlying random variables were either *i.i.d.* or Markov dependent or binary sequence of order k. Aki and Hirano (1995) derived the joint distributions of number of failures, successes and runs of success under the same set up. Under different types of counting schemes like runs of length k_1, non overlapping runs of length k_1, overlapping runs of length k_1 *etc.*, Hirano *et. al.* (1997) gave interesting results on the distributions of number of success runs of length l until the first occurrence of the success run of length k for an m^{th} order homogeneous Markov chain. The joint distributions of the waiting time and the number of outcomes such as failures, successes and success runs of length less than k under the set up of an m^{th} order homogeneous Markov chain was developed by Uchida (1998) for various enumeration schemes of runs. Chadjiconstantindis and Koutras (2001) also obtained the distribution of failures and successes in a waiting time problem.

Another scheme of μ-overlapping counting was introduced by Aki and Hirano (2000) where an overlap of at most μ successes was allowed between two consecutive runs of length k where $0 \leq \mu \leq k-1$. They also introduced the generalized Binomial distribution of order k and investigated some of it's properties. It is easy to observe that when $\mu = 0$, the counting scheme matches with the non-overlapping counting while $\mu = k-1$ yields the overlapping counting scheme. Han and Aki (2000) have extended this counting scheme for the negative values of μ in which there should be at least $|\mu|$ trials between any two consecutive success runs of length k. They have derived recurrence relations for the probability generating function (*pgf*) of the number of μ-overlapping success runs of length k. Inoue and Aki (2003) derived exact formulae for the *pgf* of the above-mentioned random variable in the case of two-state Markov dependent trials. They also derived explicitly, in the same case, the *pgf* of the waiting time until the r^{th} occurrence of the μ-overlapping success run of length k. Makri and Philippou (2005) obtained the exact formulas for the probability distribution function of the number of μ-overlapping success runs of length k in n trials. Makri *et. al.* (2007) considered the concept of μ-overlapping success runs in the Polya-Eggenberger sampling scheme and obtained the distribution of the number of drawings according to the Polya-Eggenberger sampling scheme until the r^{th} occurrence of the μ-overlapping success run of length k. They have also introduced Polya, inverse Polya, and circular Polya distributions of order k for μ-overlapping success runs of length k.

Anuradha (2023) introduced the l-look-back counting scheme for runs of successes. In this scheme, if a run has been counted starting at time i, *i.e.*, $\{X_i = X_{i+1} = \cdots = X_{i+k-1} = \mathbf{1}\}$, then no runs can be counted till the time point $i+l$ and the next counting of runs can start only from the time point $i+l+1$, where $X_i = \mathbf{1}$ represents a success at time i and l is a non-negative integer. This process is repeated every time a run is counted. In other words, if a run is counted starting at time i, then there are k-consecutive successes starting from the time point i and no runs of length k has been counted starting at time points $i-1, i-2, \ldots, i-l$. The mathematical definition has been provided in the section 3.

The look-back counting scheme generalizes the concept of run counting and encompasses both the definitions of overlapping counting as well as the non-overlapping counting thereby giving rise to new objects for further study. Indeed, if $l = 0$, this matches exactly with the counting of overlapping runs of length k, and if $l = k - 1$, this counting scheme results in the counting of non-overlapping runs of length k. It should further be noted that μ-overlapping scheme, for positive values of μ, can also be identified as l-look-back counting where $\mu = k - l - 1$. However, for negative values of μ, the definitions do not match with the corresponding value of $l = k - \mu - 1$. We illustrate this difference with the same example as cited in Han and Aki (2000). Consider the following sequence of successes and failures:

$$\mathbf{11110110001111111110000111}.$$

In this sequence, for $k = 3$ and $l = 3$, we have four 3-look-back runs of length 3 starting at trials $1, 11, 15$ and 23, while there are only three (-1)-overlapping runs of length 3, starting at $1, 11$ and 15. Therefore l-look-back counting scheme is an entirely new scheme of counting which has not yet been studied in detail.

Under the set up of m^{th} order homogeneous Markov chain, Anuradha (2023) proved that the waiting time distribution of the n^{th} occurrence of the l-look-back run of length k converges to an extended Poisson distribution when the system exhibits strong propensity towards success. Further central limit theorem was established for the number of l-look-back runs of length k till the n^{th} trial.

Aki and Hirano (2000) established that the number of $(l - 1)$-overlapping runs of length k $(l < k)$ until the n^{th} overlapping occurrence of success run of length l follows a generalized Binomial distribution of order $(k - l)$ for the *i.i.d.* as well as the m^{th} order homogeneous Markov chain. In this paper, we pose a different problem from the counting perspective. We fix two positive integers $k_1 \leq k$ and another integer $l \geq 0$ and count the number of overlapping runs of length k_1 until the n^{th} occurrence of l-look-back run of length k. The stopping time originates from the l-look-back counting scheme which encompasses non-overlapping, overlapping as well as μ-overlapping (for positive μ) counting schemes. We should also note that there is no restriction on l, which may equal or exceed k. Our focus is on counting of runs of smaller lengths (k_1) until a stopping rule which involves occurrences of runs of larger length (k). We obtain a decomposition of the number of runs until the stopping time into a sum of independent random variables. This, in turn, brings out a new discrete distribution of order k and also establishes new connections with the other known discrete distributions.

Koutras (1997) defined a Markov Negative Binomial distribution of order k where he studied the waiting time distributions associated with the runs of length k for a two-state Markov chain. In this paper, we introduce a new distribution which is different from the above. We denote it by *generalized Binomial type distribution*. The probability generating function of this distribution has been derived, which also shows how it generalizes the classical Binomial distribution and the classical negative Binomial distribution (refer to Definition 1). In our study, the generalized Binomial type distribution will play a central role, along with the extended negative binomial distribution of order k with parameters n and $(p_1, p_2, \ldots, p_k)$ which was introduced by Aki (1985).

Our results show that the number of overlapping runs of length k_1 up to the r^{th} occurrence of the l-look-back run of length k $(k_1 \leq k)$ can be split into a sum of r independent random variables. We further establish that except the first one, all the other random variables are identically distributed. The result has a number of interesting corollaries. For example, the results of Aki and Hirano (1994), on the number of successes until the first occurrence of the k consecutive successes for the *i.i.d.* as well as the Markov chain set-up can be derived as a corollary from our result (see Corollary 3). We also show that under the assumption of strong tendency towards failure after k consecutive successes, the number of overlapping success runs of length k_1 can be approximated by Poisson random variable translated by r (see Corollary 2).

We employ a new technique to prove our results. First we convert the m^{th} order Markov chain to a first order Markov chain which takes values in a finite set and recast our problem into this new set-up, *i.e.*, define the success / failure in the original chain in terms of the new chain and convert all relevant definitions in terms of the new chain. Thereafter, the main tool that we employ is the method of generating functions. We use the strong Markov property on this first order Markov chain to derive a recurrence relations between the probabilities. This, in turn, yields recurrence relations between the probability generating functions (*pgfs*). Finally we consider the generating function of the *pgfs*. Using the recurrence relations between the *pgfs* we obtain a linear equation involving the generating function of the *pgfs* which is used to establish its expression. Expanding this generating function of the *pgfs*, we obtain the expression for the individual *pgf*.

In the next section, we introduce the new discrete distribution, namely *generalized Binomial type distribution* and provide the probability generating function of the distribution. In section 3, we give all the definitions and state the main result and the corollaries. Section 4 is devoted to setting up the new Markov chain and recasting of the problem in terms of the new Markov chain. In Section 5, the proof of the main theorem has been established. In the final section, we provide the conclusion of the paper.

2. A new discrete distribution

In this section we introduce a discrete distribution which will be important for our work.

Definition 1: We say that a random variable W follows a *generalized Binomial type* distribution with parameters $0 < p < 1$, $n \geq 0$ and $t \geq 1$ (denoted by $GB(p,n,t)$) if

$$W = \sum_{i=1}^{n} W_i$$

where each $\{W_i : i = 1, \ldots, n\}$ is *i.i.d.* geometric random variable truncated at t with parameter p. In case $n = 0$, the sum should be understood as 0. In other words, for $i = 1, \ldots, n$,

$$P(W_i = u) = \begin{cases} qp^u & \text{if } 0 \leq u < t \\ p^t & \text{if } u = t \\ 0 & \text{otherwise.} \end{cases}$$

If $n = 1$, we will refer a $GB(p,n,t)$ random variable as a *generalized Bernoulli type* and we will denote it by a $GBer(p,t)$.

The probability generating function $\chi_{(p,n,t)}$ of the $GB(p,n,t)$ is given by

$$\chi_{(p,n,t)}(s) = \Big(q + qps + \cdots + qp^{t-1}s^{t-1} + p^t s^t\Big)^n. \tag{1}$$

Thus, the generating function of a $GBer(p,t)$ is given by

$$\chi_{(p,t)}(s) = \Big(q + qps + \cdots + qp^{t-1}s^{t-1} + p^t s^t\Big).$$

It should be noted that if $t = 1$, W follows the binomial distribution with parameters n and p. In this sense, this can be thought of as a generalization of the binomial distribution. Further, if n is fixed and $t \uparrow \infty$, then W follows the usual negative binomial distribution with parameters n and p. Also note that, if we set $p = \lambda/n$, then

$$\chi_{(p,n,t)}(s) = \left[1 - \frac{\lambda}{n}\Big(1 - s\Big) + o\Big(\frac{1}{n}\Big)\right]^n \to \exp\Big(-\lambda(1-s)\Big) \tag{2}$$

as $n \to \infty$. The limit is the probability generating function of a Poisson random variable with parameter λ. Hence, when p and n are related in such a way as above, then $GB(p,n,t)$ converges to a Poisson random variable as $n \to \infty$.

Another discrete distribution will be important for our results. Aki (1985) had defined an *extended negative binomial distribution of order t with parameters* n and $(p_1, p_2, \ldots, p_t)$ and gave the probability generating function as

$$\varphi(s; n, (p_1, p_2, \ldots, p_t)) = \left[\frac{p_1 p_2 \ldots p_t s^t}{1 - \sum_{j=1}^{t} p_1 p_2 \cdots p_{j-1} q_j s^j}\right]^n. \tag{3}$$

We will mostly consider the case when $p_1 = p_2 = \cdots = p_t = p$. Indeed, when $t = 1$, this is the usual negative binomial distribution with parameters $0 < p < 1$ and $n \geq 1$. When $n = 1$ and $p_1 = p_2 = \cdots = p_t = p$, we will call this distribution as *extended geometric distribution of order t* with parameter p.

3. Definitions and statement of results

Let $X_{-m+1}, \ldots, X_0, X_1, \ldots$ be a sequence of stationary m-order $\{\mathbf{0}, \mathbf{1}\}$ valued Markov chain. Assume that the states of $X_{-m+1}, \ldots, X_0$ are known *i.e.*, $x_0, x_{-1}, \ldots, x_{-m+1}$ are known and we take the initial state as $X_0 = x_0, X_{-1} = x_{-1}, \ldots, X_{-m+1} = x_{-m+1}$.

Define the set $S_i = \{0, 1, \ldots, 2^i - 1\}$ for any $i \geq 0$. It is clear that S_i and $\{\mathbf{0}, \mathbf{1}\}^i$ can be connected by the one-to-one and onto mapping $x = (x_0, x_1, \ldots, x_{i-1}) \longrightarrow \sum_{j=0}^{i-1} 2^j x_j$. Since $\{X_n : n \geq -m+1\}$ is the m^{th} order Markov chain, we have the transition probabilities

$$p_x = \mathbb{P}(X_{n+1} = 1 | X_n = x_0, X_{n-1} = x_1, \ldots, X_{n-m+1} = x_{m-1}) \tag{4}$$

where $x = \sum_{j=0}^{m-1} 2^j x_j \in S_m$, for any $n \geq 0$. Therefore, we have $q_x = \mathbb{P}(X_{n+1} = 0 | X_n = x_0, X_{n-1} = x_1, \ldots, X_{n-m+1} = x_{m-1}) = 1 - p_x$. We assume that $0 < p_x < 1$ for all $x \in S_m$.

Definition 2: (**l-look-back run**) (Anuradha (2023)) Fix two integers $k \geq 1$ and $l \geq 0$. We set $R_i(k,l) = 0$ for $i = 0, -1, \ldots, -l+1$ and for any $i \geq 1$, define inductively,

$$R_i(k,l) = \prod_{j=i-1}^{i-l} (1 - R_j(k,l)) \prod_{j=i}^{i+k-1} X_j \tag{5}$$

where the first product is to be taken as 1 when $l = 0$. If $R_i(k,l) = 1$, we say that a *l-look-back run of length k* has been recorded which started at time i.

It should be noted that for a l-look-back run to start at the time point i, we need to look back at the preceding l many time points, *i.e.*, $i-1$ to $i-l$, none of which can be the starting point of a l-look-back run of length k.

Next we define the stopping times where the r^{th} l-look-back run of length k is completed.

Definition 3: For $r \geq 1$, the stopping time $\tau_r(k,l)$ be the (random) time point at which the r^{th} l-look-back run of length k is completed. In other words,

$$\tau_r(k,l) = k - 1 + \inf\{n : \sum_{i=1}^{n} R_i(k,l) = r\}. \tag{6}$$

Note that r^{th} l-look-back run of length k is completed at time point $\tau_r(k,l)$. Next we define the overlapping runs of length k.

Definition 4: (**Overlapping runs of length $\boldsymbol{k}$**) When $k (\geq 1)$ consecutive successes occur, we call it an overlapping run of length k.

We may represent this mathematically as follows:

$$R_i^{(k)} = \prod_{j=1}^{k} X_{i+j-1}.$$

Note here that $R_i^{(k)} = 1$ if and only if an overlapping run of length k starts at time point i. Here a trial can contribute to more than one runs. Indeed, if $k+1$ successes appear consecutively, starting from time i, two overlapping runs will be counted with first one starting at i and the next one starting at $i+1$. Clearly all successes between time $i+1$ to $i+k-1$ will contribute to two overlapping runs.

Let $N_n(k)$ be the number of occurrences of overlapping runs of length k until time n. In other words,

$$N_n(k) = \sum_{i=1}^{n-k+1} R_i^{(k)}.$$

In this paper, we study the number of overlapping runs of length k_1 till the stopping time $\tau_r(k,l)$ (see Definition (3)). Fix any constant $k_1 \leq k$. For each $r \geq 1$, we define the random variable

$$N_r(k_1) := N_{\tau_r(k,l)}(k_1) = \sum_{i=1}^{\tau_r(k,l)} R_i^{(k_1)} \tag{7}$$

as the number of overlapping runs of length k_1 until the stopping time $\tau_r(k,l)$.

Let us consider the following example to facilitate understanding: Consider the following sequence of **1**'s and **0**'s of length 25

$$\mathbf{1110111010110111111101011}.$$

Let $k = 3$ and $l = 1$. Now using the definition we have $R_1(3,1) = R_5(3,1) = R_{14}(3,1) = R_{16}(3,1) = R_{18}(3,1) = 1$, while for other values of i, $R_i(3,1) = 0$. Thus, stopping times become $\tau_1(3,1) = 3, \tau_2(3,1) = 7, \tau_3(3,1) = 16, \tau_4(3,1) = 18$ and $\tau_5(3,1) = 20$. For $k_1 = 2$, the number of the overlapping runs of length 2 till the stopping times are given by $N_1(2) = 2, N_2(2) = 4, N_3(2) = 7$ and $N_4(2) = 9$ and $N_5(2) = 11$.

Let us define the probability generating function of $N_r(k_1)$ as follows

$$\zeta_r(s;k_1) = \sum_{n=0}^{\infty} \mathbb{P}(N_r(k_1) = n)s^n = \sum_{n=0}^{\infty} g_r(n;k_1)s^n. \tag{8}$$

Now we state our main result which we prove in Section 5.

Theorem 1: For any initial condition $x \in S_m$ and $k_2 = k - k_1$ and $k_1 \geq m$, the probability generating function of $N_r(k_1)$ is given by,

$$\begin{aligned}\zeta_r(s;k_1) = &\frac{s\Big(p_{2^m-1}s\Big)^{k_2}}{1 - \sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}}\Bigg[\Big(p_{2^m-1}s\Big)^{l+1} \\ &+ \frac{s\Big(p_{2^m-1}s\Big)^{k_2}}{1 - \sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}} \sum_{j=0}^{l} q_{2^m-1}\Big(p_{2^m-1}s\Big)^j\Bigg]^{r-1}.\end{aligned}$$

In the above and subsequently, we have used the convention that the sum is taken to be 0 if the starting index of the sum is bigger than the ending index of the sum (which happens in the above expression when we take $k_2 = 0$).

Now the result of theorem 1 provides a powerful representation of $N_r(k_1)$ through the extended geometric random variables and generalized Bernoulli type distribution.

Let us define the indicator function as follows:

$$\mathbb{I}_{\{u\}}(v) = \begin{cases} 1 & \text{if } u = v \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Corollary 1: Suppose that $\{G_i^{(E)} : i = 1, \dots, r\}$ and $\{B_i^{(G)} : i = 1, \dots, r\}$ are independent families of *i.i.d.* random variables where each $G_i^{(E)}$ is having an extended geometric distribution of order k_2 with parameter p_{2^m-1} and each $B_i^{(G)}$ is a generalized Bernoulli type random variable $\mathrm{GBer}(p_{2^m-1}, l+1)$. Then

$$N_r(k_1) \stackrel{d}{=} \Big(1 + G_1^{(E)}\Big) + \sum_{i=2}^{r}\Big[B_i^{(G)} + \Big(1 + G_i^{(E)}\Big)\Big(1 - \mathbb{I}_{\{l+1\}}\Big(B_i^{(G)}\Big)\Big)\Big].$$

Indeed, we have that the generating function of any $G_i^{(E)}$ is given by the equation (3). Also, the generating function of $B_i^{(G)} + \left(1 + G_i^{(E)}\right)\left(1 - \mathbb{I}_{\{l+1\}}\left(B_i^{(G)}\right)\right)$ is given by

$$\begin{aligned}
&\sum_{i=0}^{\infty}\sum_{j=0}^{l+1} s^{j+(1+i)\left(1-\mathbb{I}_{\{l+1\}}(j)\right)}\mathbb{P}(G_i^{(E)}=i)\mathbb{P}(B_i^{(G)}=j)\\
&=\sum_{i=0}^{\infty}\sum_{j=0}^{l} s^{j+(1+i)}\mathbb{P}(G_i^{(E)}=i)\mathbb{P}(B_i^{(G)}=j)+\sum_{i=0}^{\infty} s^{l+1}\mathbb{P}(G_i^{(E)}=i)\mathbb{P}(B_i^{(G)}=l+1)\\
&=s\sum_{i=0}^{\infty} s^i\mathbb{P}(G_i^{(E)}=i)\sum_{j=0}^{l} s^j\mathbb{P}(B_i^{(G)}=j)+s^{l+1}\left(p_{2^m-1}\right)^{l+1}\sum_{i=0}^{\infty}\mathbb{P}(G_i^{(E)}=i)\\
&=\frac{s\left(p_{2^m-1}s\right)^{k_2}}{1-\sum_{j=0}^{k_2-1} q_{2^m-1}\left(p_{2^m-1}\right)^j s^{j+1}}\sum_{j=0}^{l} q_{2^m-1}\left(p_{2^m-1}s\right)^j+\left(p_{2^m-1}s\right)^{l+1}.
\end{aligned}$$

Thus using the independence of the random variables, we now conclude that the generating functions of the random variables of both sides of the corollary 1 are same. This proves the corollary.

If $r = 1$, the distribution of $N_1(k_1)$ is actually an extended geometric distribution of order k_2 and parameter p_{2^m-1} translated by 1. If $k_2 = 0$, *i.e.*, $k = k_1$, we have that $G_i^{(E)} = 0$ and hence we have

$$\begin{aligned}
N_r(k_1) &\overset{d}{=} 1+\sum_{i=2}^{r}\left[B_i^{(G)}+1-\mathbb{I}_{\{l+1\}}\left(B_i^{(G)}\right)\right]\\
&= r+\sum_{i=2}^{r}\left[B_i^{(G)}-\mathbb{I}_{\{l+1\}}\left(B_i^{(G)}\right)\right]=r+\sum_{i=1}^{r-1} D_i^{(G)}
\end{aligned}$$

where $D_i^{(G)} = B_{i+1}^{(G)} - \mathbb{I}_{\{l+1\}}\left(B_{i+1}^{(G)}\right)$ for $i = 1, 2, \ldots, r-1$. Now, we observe that $D_i^{(G)}$ has a geometric distribution truncated at l. Indeed, for $j < l$, it is easy to see that $\mathbb{P}(D_i^{(G)} = j) = \mathbb{P}(B_{i+1}^{(G)} = j) = q_{2^m-1}\left(p_{2^m-1}\right)^j$ and for $j = l$, we have $\mathbb{P}(D_i^{(G)} = l) = \mathbb{P}(B_{i+1}^{(G)} = l) + \mathbb{P}(B_{i+1}^{(G)} = l+1) = q_{2^m-1}\left(p_{2^m-1}\right)^l + \left(p_{2^m-1}\right)^{l+1} = \left(p_{2^m-1}\right)^l$. Thus, $N_r(k_1) - r$ has generalized Binomial type distribution with parameters $p_{2^m-1}, r-1$ and l.

Under the assumption that the system has a strong tendency towards failure when m consecutive successes are observed, *i.e.*, p_{2^m-1} as a function of r converges to 0 in such a way that

$$rp_{2^m-1} \to \lambda > 0 \text{ as } r \to \infty, \tag{10}$$

using the equation (2) and the subsequent discussion, we can easily obtain the following corollary.

Corollary 2: For any initial condition $x \in S_m$, if the condition (10) holds and if $k_2 = 0$, we have

$$N_r(k_1) - r \quad \Rightarrow \mathrm{Poi}(\lambda)$$

where $\mathrm{Poi}(\lambda)$ is the Poisson distribution with parameter λ.

If we set $k_1 = 1$, $N_r(k_1)$ represents the number of successes till the r^{th} occurrence of the l-look-back run of length k. Thus, we have the following corollary:

Corollary 3: For the *i.i.d.* case or the Markov dependent case, the probability generating function of the number of successes till the r^{th} occurrence of the l-look-back run of length k, *i.e.*, $N_r(1)$ is given by,

$$\zeta_r(s;1) = \frac{s\Big(p_{2^m-1}s\Big)^{k-1}}{1-\sum_{j=0}^{k-2} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}}\Bigg[\Big(p_{2^m-1}s\Big)^{l+1} + \frac{s\Big(p_{2^m-1}s\Big)^{k-1}}{1-\sum_{j=0}^{k-2} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}}\sum_{j=0}^{l} q_{2^m-1}\Big(p_{2^m-1}s\Big)^j\Bigg]^{r-1}.$$

For $r = 1$, the expression reduces to

$$\zeta_1(s;1) = \frac{\Big(p_{2^m-1}\Big)^{k-1}(1-p_{2^m-1}s)s^k}{1-s+q_{2^m-1}\Big(p_{2^m-1}\Big)^{k-1}s^k}$$

which is the probability generating function of the number of successes until the first occurrence of k consecutive successes. For the *i.i.d.* case, we have $p_{2^m-1} = p$ and for the Markov dependent case, we have $p_{2^m-1} = p_{11}$. Putting these values, we observe that we may obtain the results (Proposition 3.4 and Theorem 3.2) of Aki and Hirano (1994). Therefore our result provides a generalized version of *pgf* for all values of r.

4. A new Markov chain

Now we outline the underlying set up which will be used in the subsequent sections to establish the results. Let us define two functions $f_0, f_1 : S_{k_1} \to S_{k_1}$ by

$$f_1(x) = 2x + 1 \pmod{2^{k_1}} \text{ and } f_0(x) = 2x \pmod{2^{k_1}}.$$

Further define a projection $\theta_m : S_{k_1} \to S_m$ by $\theta_m(x) = x \pmod{2^m}$. Now, set $X_{-m} = X_{-m-1} = \cdots = X_{-k_1+1} = 0$. Define a sequence of random variables $\{Y_n : n \geq 0\}$ as follows:

$$Y_n = \sum_{j=0}^{k_1-1} 2^j X_{n-j}.$$

Since $X_i \in \{\mathbf{0}, \mathbf{1}\}$ for all i, Y_n assumes values in the set S_{k_1}. The random variables X_n's are stationary and forms an m^{th} order Markov chain, hence we have that $\{Y_n : n \geq 0\}$ is a homogeneous Markov chain with transition matrix given by

$$\mathbb{P}(Y_{n+1} = y | Y_n = x) = \begin{cases} p_{\theta_m(x)} & \text{if } y = f_1(x) \\ 1 - p_{\theta_m(x)} & \text{if } y = f_0(x) \\ 0 & \text{otherwise.} \end{cases}$$

Note that Y_n is even if and only if $X_n = 0$. This motivates us to define the function $\kappa : S_{k_1} \to \{0, 1\}$ by

$$\kappa(x) = \begin{cases} 1 & \text{if } x \text{ is odd} \\ 0 & \text{if } x \text{ is even.} \end{cases}$$

Therefore, $\kappa(Y_n) = 1$ if and only if $X_n = 1$. Hence, the definition of l-look-back run can be described in terms of Y_n's as

$$R_i(k,l) = \prod_{j=i-l}^{i-1} (1 - R_j(k,l)) \prod_{j=i}^{i+k-1} \kappa(Y_j).$$

Let us fix any initial condition $x \in S_m$. We denote the probability measure governing the distribution of $\{Y_n : n \geq 1\}$ with $Y_0 = x \in S_k$ by $\mathbb{P}_x$. Since we have set $X_{-m} = X_{-m-1} = \cdots = X_{-k+1} = 0$, we have $Y_0 = x$.

In order to obtain the recurrence relation for the probabilities, we will condition the process after the first occurrence of the run of length k_1. Therefore, we consider the stopping time T when the first occurrence of a run of length k_1 ends, *i.e.*, when we observe k_1 successes consecutively for the first time. More precisely, define

$$T := \inf\{i \geq k_1 : \prod_{j=i-k_1+1}^{i} X_j = 1\}. \tag{11}$$

We would like to translate the above definition to Y_i's. It must be the case that when T occurs, last k_1 trials have resulted in success, which may be described by $\kappa(Y_j) = 1$ for $j = i - k_1 + 1$ to i. Therefore, Y_T must equal $2^{k_1} - 1$. Since this is the first occurrence, this has not happened earlier. So, T can be better described as

$$T = \inf\{i \geq k_1 : Y_i = 2^{k_1} - 1\},$$

i.e., the first visit of the chain to the state $2^{k_1} - 1$ after time $k_1 - 1$. Now, we note that $\{Y_n : n \geq 0\}$ is a Markov chain with finite state space. Further, since $0 < p_u < 1$ for $u \in S_m$, this is an irreducible chain; hence, it is positive recurrent. So we must have $\mathbb{P}_x(T < \infty) = 1$. We observe that when the first occurrence of k consecutive successes happen, we must have the occurrence of k_1 successes previously since $k_1 \leq k$. Therefore, we have $\mathbb{P}_x(T \leq \tau_1(k,l)) = 1$.

5. Overlapping runs till the stopping time

In this section, we study the distribution of overlapping runs of length k_1. We will employ the method of generating functions to derive these results.We obtain a recurrence relation between the probabilities in order to derive the generating functions.

Let us define the probability, for $x \in S_m, n \in \mathbb{Z}$,

$$g_r^{(x)}(n; k_1) = \mathbb{P}_x\Big(N_r(k_1) = n\Big). \tag{12}$$

We note that since $N_r(k_1) \geq 1$, $\mathbb{P}_x\Big(N_r(k_1) = n\Big) = 0$ for $n \leq 0$. Also, if $r = 1$ and $k_2 = k - k_1 = 0$, i.e, $k = k_1$, we have that $N_1(k_1) = 1$.

We will show that these probabilities $g_r^{(x)}(n;k_1)$ is actually independent of the initial condition x. First we consider the case when $r=1$. As we have already observed, if $k_2=0$,

$$g_1^{(x)}(1;k_1)=\mathbb{I}_{\{1\}}(n)$$

where $\mathbb{I}_{\{u\}}(v)$ is the indicator function defined in (9). Clearly we have $g_1^{(x)}(n;k_1)$ is independent of x.

Now, we concentrate on the case when $r=1$ and $k_2=k-k_1>0$, *i.e.*, $k>k_1$. We note that $N_1(k_1)\geq(k_2+1)$ and hence $\mathbb{P}_x(N_1(k_1)=n)=g_1^{(x)}(n;k_1)=0$ for $n\leq k_2$.

Theorem 2: For $n>k_2$ and $k_2=k-k_1>0$, we have

$$g_1^{(x)}(n;k_1)=\sum_{t=0}^{k_2-1}q_{2^m-1}\Big(p_{2^m-1}\Big)^t g_1^{(2^m-2)}(n-t-1;k_1)+\Big(p_{2^m-1}\Big)^{k_2}\mathbb{I}_{\{k_2+1\}}(n) \tag{13}$$

where $\mathbb{I}_{\{u_1\}}(u_2)$ is the indicator function defined in (9).

Proof: When $k_2=k-k_1>0$ and $r=1$, using the fact that $Y_T=2^{k_1}-1$ with probability 1, we have

$$\begin{aligned}
&g_1^{(x)}(n;k_1)=\mathbb{P}_x(N_1(k_1)=n)=\mathbb{P}_x(N_1(k_1)=n,Y_T=2^{k_1}-1)\\
&=\sum_{t=0}^{k_2-1}\mathbb{P}_x(N_1(k_1)=n,Y_T=2^{k_1}-1,Y_{T+1}=2^{k_1}-1,\dots,\\
&\qquad Y_{T+t}=2^{k_1}-1,Y_{T+t+1}=2^{k_1}-2)\\
&+\mathbb{P}_x(N_1(k_1)=n,Y_T=2^{k_1}-1Y_{T+1}=2^{k_1}-1,Y_{T+2}=2^{k_1}-1,\dots,\\
&\qquad Y_{T+k_2-1}=2^{k_1}-1,Y_{T+k_2}=2^{k_1}-1).
\end{aligned}\tag{14}$$

We look at the terms in the summation first. For any $0\leq t\leq k_2-1$, we have,

$$\begin{aligned}
&\mathbb{P}_x(N_1(k_1)=n,Y_T=2^{k_1}-1,Y_{T+1}=2^{k_1}-1,\dots,Y_{T+t}=2^{k_1}-1,Y_{T+t+1}=2^{k_1}-2)\\
&=\mathbb{P}_x(N_1(k_1)=n\mid Y_T=2^{k_1}-1,Y_{T+1}=2^{k_1}-1,\dots,Y_{T+t}=2^{k_1}-1,Y_{T+t+1}=2^{k_1}-2)\\
&\times\mathbb{P}_x(Y_T=2^{k_1}-1,Y_{T+1}=2^{k_1}-1,\dots,Y_{T+t}=2^{k_1}-1,Y_{T+t+1}=2^{k_1}-2).
\end{aligned}\tag{15}$$

The second term in (15) can be written as

$$\begin{aligned}
&\mathbb{P}_x(Y_T=2^{k_1}-1,Y_{T+1}=2^{k_1}-1,\dots,Y_{T+t}=2^{k_1}-1,Y_{T+t+1}=2^{k_1}-2)\\
&=\mathbb{P}_x(Y_{T+t+1}=2^{k_1}-2\mid Y_T=2^{k_1}-1,Y_{T+1}=2^{k_1}-1,\dots,Y_{T+t}=2^{k_1}-1)\\
&\times\prod_{j=1}^{t}\mathbb{P}_x(Y_{T+j}=2^{k_1}-1\mid Y_T=2^{k_1}-1,Y_{T+1}=2^{k_1}-1,\dots,Y_{T+j-1}=2^{k_1}-1).
\end{aligned}$$

Now, $T+j-1$ is also a stopping time for any $1\leq j\leq t$. We denote by $\mathcal{F}_{T+j-1}$, the σ-algebra generated by the process Y_n up to the stopping time $T+j-1$, and by $\mathcal{F}_{(T+j-1)+}$, the σ-algebra generated by the process after the stopping time $T+j-1$. Clearly,

$\{Y_{T+1} = 2^{k_1} - 1, \dots, Y_{T+j-1} = 2^{k_1} - 1\} \in \mathcal{F}_{T+j-1}$ and $\{Y_{T+j} = 2^{k_1} - 1\} \in \mathcal{F}_{(T+j-1)+}$. Thus, by strong Markov property, we can write

$$\begin{aligned}
&\mathbb{P}_x(Y_{T+j} = 2^{k_1} - 1 \mid Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \dots, Y_{T+j-1} = 2^{k_1} - 1) \\
&= \mathbb{P}_{Y_{T+j-1}}(Y_{T+j} = 2^{k_1} - 1) = \mathbb{P}_{2^{k_1}-1}(Y_1 = 2^{k_1} - 1) = p_{2^m-1}. \qquad (16)
\end{aligned}$$

A similar argument shows that

$$\mathbb{P}_x(Y_{T+t+1} = 2^{k_1} - 2 \mid Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \dots, Y_{T+t} = 2^{k_1} - 1) = q_{2^m-1}. \qquad (17)$$

For the first term in (15), we note that $T + t + 1$ is also a stopping time and $\{Y_{T+1} = 2^{k_1} - 1, \dots, Y_{T+t} = 2^{k_1} - 1, Y_{T+t+1} = 2^{k_1} - 2\} \in \mathcal{F}_{T+t+1}$. Since $Y_{\tau_1(k_1)} = 2^{k_1} - 1$, we must have $X_{T-k_1} = 0$ and $X_{T-j} = 1$ for $j = 0, 1, \dots, k_1 - 1$. Further, since $Y_{\tau_1(k_1)+j} = 2^{k_1} - 1$ for $j = 1, \dots, t$ and $Y_{T+t+1} = 2^{k_1} - 2$, we also have $X_{T+j} = 1$ for $j = 0, 1, \dots, t$ and $X_{T+t+1} = 0$. Therefore, we have a sequence of $\mathbf{1}$'s of length $k_1 + t$ with $t > 0$ which contributes to $t + 1$ overlapping runs of length k_1 and since there are no runs of length k_1 before T, by the very definition of T, we have that the number of overlapping runs of length k_1 up to time $T+t+1$ is $1 + t$. Since $t \leq k_2 - 1$, we have that $T + t + 1 < \tau_1(k, l)$. Let us define $Y'_i = Y_{i+T+t+1}$ for $i \geq 0$. Now, using the strong Markov property, we have that $\{Y'_i : i \geq 0\}$ is a homogeneous Markov chain with same transition matrix as that of $\{Y_i : i \geq 0\}$ with $Y'_0 = 2^{k_1} - 2$. Now, define $\tau'_1(k, l)$ as the stopping time for the process $\{Y'_i : i \geq 0\}$. From the above discussion, we have that $\tau_1(k, l) = T + t + 1 + \tau'_1(k, l)$. Further, if we define, $N'_1(k_1)$ as the number of overlapping runs of length k_1 up to time $\tau'_1(k, l)$ for the process $\{Y'_i : i \geq 0\}$, we must have that $N'_1(k_1) = n - t - 1$. Therefore, we have,

$$\begin{aligned}
&\mathbb{P}_x(N_1(k_1) = n \mid Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \dots, Y_{T+t} = 2^{k_1} - 1, Y_{T+t+1} = 2^{k_1} - 2) \\
&= \mathbb{P}_{(2^m-2)}(N'_1(k_1) = n - t - 1) = g_1^{(2^m-2)}(n - t - 1; k_1). \qquad (18)
\end{aligned}$$

The last term in (14) can be similarly written as

$$\begin{aligned}
&\mathbb{P}_x(N_1(k_1) = n, Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \dots, Y_{T+k_2} = 2^{k_1} - 1) \\
&= \prod_{j=1}^{k_2} \mathbb{P}_x(Y_{T+j} = 2^{k_1} - 1 \mid Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \dots, Y_{T+j-1} = 2^{k_1} - 1) \\
&\quad \times \mathbb{P}_x(N_1(k_1) = n \mid Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \dots, \\
&\qquad\qquad Y_{T+k_2-1} = 2^{k_1} - 1, Y_{T+k_2} = 2^{k_1} - 1) \\
&= \left(p_{2^m-1}\right)^{k_2} \mathbb{P}_x(N_1(k_1) = n \mid Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \dots, \\
&\qquad\qquad\qquad Y_{T+k_2-1} = 2^{k_1} - 1, Y_{T+k_2} = 2^{k_1} - 1).
\end{aligned}$$

Note that given $\{Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \dots, Y_{T+k_2-1} = 2^{k_1} - 1, Y_{T+k_2} = 2^{k_1} - 1\}$, we have $\tau_1(k, l) = T + k_2$. Further, in such a case we have exactly $k_2 + 1$ many overlapping runs of length k_1 until time $T + k_2$. Therefore, $N_1(k_1) = n$ if and only if $n = k_2 + 1$. In other words, $\mathbb{P}_x(N_1(k_1) = n \mid Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \dots, Y_{T+k_2-1} = 2^{k_1} - 1, Y_{T+k_2} = 2^{k_1} - 1) = 1_{\{k_2+1\}}(n)$ where $\mathbb{I}$ is the indicator function as defined in (9).

Thus combining the above equation with equations (14) - (18), we can express

$$g_1^{(x)}(n;k_1) = \sum_{t=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^t g_1^{(2^m-2)}(n-t-1;k_1) + \Big(p_{2^m-1}\Big)^{k_2}\mathbb{I}_{\{k_2+1\}}(n).$$

This completes the proof. $\square$

We note that the right hand side of (13) does not involve the initial condition $x \in S_m$. Therefore $g_1^{(x)}(n;k_1)$ must be independent of x. So, we will drop x and denote the above probability by $g_1(n;k_1)$. Thus, we have the following corollary from theorem 2.

Corollary 4: For $n \geq k_2+1$ and $k_2 > 0$, we have

$$g_1(n;k_1) = \sum_{t=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^t g_1(n-t-1;k_1) + \Big(p_{2^m-1}\Big)^{k_2}\mathbb{I}_{\{k_2+1\}}(n). \tag{19}$$

Let us recall that

$$\zeta_r(s;k_1) = \sum_{n=0}^{\infty} \mathbb{P}(N_r(k_1)=n)s^n = \sum_{n=0}^{\infty} g_r(n;k_1)s^n.$$

When $k_2 = 0$, we have

$$\zeta_1(s;k_1) = s.$$

For $k_2 > 0$, we may use the equation (19) to derive its generating function. We have

$$\begin{aligned}
\zeta_1(s;k_1) &= \sum_{n=k_2+1}^{\infty} g_r(n;k_1)s^n \\
&= \sum_{n=k_2+1}^{\infty}\Big[\sum_{t=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^t g_1(n-t-1;k_1) + \Big(p_{2^m-1}\Big)^{k_2}\mathbb{I}_{\{k_2+1\}}(n)\Big]s^n \\
&= \Big(p_{2^m-1}\Big)^{k_2}s^{k_2+1} + \sum_{t=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^t s^{t+1}\sum_{n=k_2+1}^{\infty} g_1(n-t-1;k_1)s^{n-t-1} \\
&= \Big(p_{2^m-1}\Big)^{k_2}s^{k_2+1} + \zeta_1(s;k_1)\sum_{t=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^t s^{t+1}.
\end{aligned}$$

This linear equation can now be solved to yield the following corollary.

Corollary 5: For $r = 1$, we have

$$\zeta_1(s;k_1) = \frac{s\Big(p_{2^m-1}s\Big)^{k_2}}{1-\sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}}. \tag{20}$$

Now we consider the case when $r > 1$. In this case also, we note that $N_r(k_1) \geq (k_2+1)+(r-1)(l+1)$. Hence $g_r^{(x)}(n;k_1) = \mathbb{P}_x(N_r(k_1)=n) = 0$ for $n \leq (r-1)(l+1)+k_2$. Now, we derive the recurrence relation.

Theorem 3: For $n \geq (k_2+1)+(r-1)(l+1)$ and $x \in S_m$, we have

$$\begin{aligned} g_r^{(x)}(n;k_1) &= \Big(p_{2^m-1}\Big)^{k_2+(r-1)(l+1)} \mathbb{I}_{\{n\}}(k_2+(r-1)(l+1)+1) \\ &+ \sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j g_r^{(2^m-2)}(n-j-1;k_1) \\ &+ \sum_{j_1=0}^{r-2}\sum_{j_2=0}^{l} q_{2^m-1}\Big(p_{2^m-1}\Big)^{k_2+j_1(l+1)+j_2} g_{r-1-j_1}^{(2^m-2)}(n-1-k_2-j_1(l+1)-j_2;k_1). \end{aligned} \tag{21}$$

where $\mathbb{I}_{v_1}(v_2)$ is the indicator function, as defined in the previous theorem.

Proof: We proceed in the same way as in the previous theorem. Conditioning on the first occurrence of k_1 many successes, *i.e.*, T, we have, for any $n \geq (k_2+1)+(r-1)(l+1)$,

$$\begin{aligned} g_r^{(x)}(n;k_1) = &\sum_{t=0}^{k_2+(r-1)(l+1)-1} \mathbb{P}_x(N_r(k_1)=n, Y_T=2^{k_1}-1, Y_{T+1}=2^{k_1}-1,\ldots, \\ &\qquad Y_{T+t}=2^{k_1}-1, Y_{T+t+1}=2^{k_1}-2) \\ &+ \mathbb{P}_x(N_r(k_1)=n, Y_T=2^{k_1}-1, Y_{T+1}=2^{k_1}-1, Y_{T+2}=2^{k_1}-1,\ldots, \\ &\qquad Y_{T+(r-1)(l+1)-1}=2^{k_1}-1, Y_{T+(r-1)(l+1)}=2^{k_1}-1). \end{aligned} \tag{22}$$

The last term in (22) is similar to the last term in equation (14) in the previous theorem. Thus this term can be simplified in the similar way. Indeed using the same arguments, as done after equation (18), we get

$$\begin{aligned} &\mathbb{P}_x(N_r(k_1)=n, Y_T=2^{k_1}-1, Y_{T+1}=2^{k_1}-1, Y_{T+2}=2^{k_1}-1,\ldots, \\ &\qquad Y_{T+(r-1)(l+1)-1}=2^{k_1}-1, Y_{T+(r-1)(l+1)}=2^{k_1}-1) \\ &= \Big(p_{2^m-1}\Big)^{k_2+(r-1)(l+1)} \mathbb{I}_{\{n\}}(k_2+(r-1)(l+1)+1). \end{aligned} \tag{23}$$

The terms in the summation in (22) can also be handled in the similar way as done in the previous theorem. Fix any j with $0 \leq t \leq k_2+(r-1)(l+1)-1$ and we obtain that

$$\begin{aligned} &\mathbb{P}_x(N_r(k_1)=n, Y_T=2^{k_1}-1, Y_{T+1}=2^{k_1}-1,\ldots, Y_{T+t}=2^{k_1}-1, Y_{T+t+1}=2^{k_1}-2) \\ &= \mathbb{P}_x(N_r(k_1)=n \mid Y_T=2^{k_1}-1, Y_{T+1}=2^{k_1}-1,\ldots, Y_{T+t}=2^{k_1}-1, Y_{T+t+1}=2^{k_1}-2) \\ &\times \mathbb{P}(Y_T=2^{k_1}-1, Y_{T+1}=2^{k_1}-1,\ldots, Y_{T+t}=2^{k_1}-1, Y_{T+t+1}=2^{k_1}-2). \end{aligned} \tag{24}$$

The last term in the product above is again simplified using the product of conditional terms and the strong Markov property. Since $Y_T = 2^{k_1} - 1$ with probability 1, we have

$$\begin{aligned}
&\mathbb{P}(Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \ldots, Y_{T+t} = 2^{k_1} - 1, Y_{T+t+1} = 2^{k_1} - 2) \\
&= \mathbb{P}_x(Y_{T+t+1} = 2^{k_1} - 2 \mid Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \ldots, Y_{T+t} = 2^{k_1} - 1) \\
&\times \prod_{j=1}^{t} \mathbb{P}_x(Y_{T+j} = 2^{k_1} - 1 \mid Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \ldots, Y_{T+j-1} = 2^{k_1} - 1) \\
&= \mathbb{P}_x(Y_{T+t+1} = 2^{k_1} - 2 \mid Y_{T+t} = 2^{k_1} - 1) \times \prod_{j=1}^{t} \mathbb{P}_x(Y_{T+j} = 2^{k_1} - 1 \mid Y_{T+j-1} = 2^{k_1} - 1) \\
&= q_{2^m-1}\Big(p_{2^m-1}\Big)^t. \qquad (25)
\end{aligned}$$

For the first term, we note that the event $\{Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \ldots, Y_{T+t} = 2^{k_1} - 1, Y_{T+t+1} = 2^{k_1} - 2\}$ implies that at time $T + t + 1$, we have just observed $k_1 + t$ many successes followed by a failure. This string of $k_1 + t$ many successes, will contribute $t + 1$ many overlapping runs of successes. Since T is the first time when we observe first k_1 many consecutive successes, we have $t + 1$ overlapping success runs completed at time $T + t + 1$. Thus, we are left with $n - t - 1$ many runs for the remaining part, *i.e.*, after time $T + t + 1$.

At time $T + t + 1$, we have the information that $k_1 + t$ many successes followed by a failure has just been observed. Using the strong Markov property, we can think that the process restarts with this information. In other words, considering the converted Y process, we are restarting the process with the initial condition $Y_{T+t+1} = 2^{k_1} - 2$.

Now, we examine the two cases namely $k_2 = k - k_1 = 0$, *i.e.*, $k = k_1$ and $k_2 > 0$, *i.e.*, $k > k_1$ separately. When $k_2 = 0$ and $k_1 + t$ many successes followed by a failure has just been observed, then we have already completed $1 + \lfloor t/(l + 1)\rfloor$ many l-look-back runs of length k where $\lfloor a\rfloor$ is the largest integer smaller or equal to a. Thus, we are left with $r - \lfloor t/(l+1)\rfloor - 1$ many l-look-back runs of length k, which is to be completed by the process after the time $T + t + 1$. Hence, we obtain,

$$\begin{aligned}
&\mathbb{P}_x(N_r(k_1) = n \mid Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \ldots, Y_{T+t} = 2^{k_1} - 1, Y_{T+t+1} = 2^{k_1} - 2) \\
&= \mathbb{P}_{2^m-2}(N_{r-\lfloor t/(l+1)\rfloor-1}(k_1) = (n - t - 1)) = g^{(2^m-2)}_{r-\lfloor t/(l+1)\rfloor-1}(n - t - 1; k_1). \qquad (26)
\end{aligned}$$

For $k_2 > 0$, the argument is essentially the same, except for one part. When $t \le k_2-1$, we would have $k_1 + t \le k_1 + k - k_1 - 1 = k - 1$ many successes followed by a failure. This will not contribute to any run of l-look-back run of length k. But for $t \ge k_2$, we will have $1 + \lfloor (t - k_2)/(l + 1)\rfloor$ many l-look-back runs of length k which have been completed. Thus, we have

$$\begin{aligned}
&\mathbb{P}_x(N_r(k_1) = n \mid Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, \ldots, Y_{T+t} = 2^{k_1} - 1, Y_{T+t+1} = 2^{k_1} - 2) \\
&= \begin{cases} g^{(2^m-2)}_r(n - t - 1; k_1) & \text{if } t \le k_2 - 1 \\ g^{(2^m-2)}_{r-\lfloor (t-k_2)/(l+1)\rfloor-1}(n - t - 1; k_1) & \text{if } t \ge k_2. \end{cases} \qquad (27)
\end{aligned}$$

Therefore, combining all the terms above from equations (24), (25), (26) and (27), we have

$$\begin{aligned}
&\sum_{t=0}^{k_2+(r-1)(l+1)-1} \mathbb{P}_x(N_r(k_1) = n, Y_T = 2^{k_1} - 1, Y_{T+1} = 2^{k_1} - 1, Y_{T+2} = 2^{k_1} - 1, \dots, \\
&\qquad\qquad Y_{T+t} = 2^{k_1} - 1, Y_{T+t+1} = 2^{k_1} - 2) \\
&= \sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j g_r^{(2^m-2)}(n-j-1;k_1) \\
&\quad + \sum_{j_1=0}^{r-2}\sum_{j_2=0}^{l} q_{2^m-1}\Big(p_{2^m-1}\Big)^{k_2+j_1(l+1)+j_2} g_{r-1-j_1}^{(2^m-2)}(n-1-k_2-j_1(l+1)-j_2;k_1). \qquad (28)
\end{aligned}$$

Now combining the equations (22), (23) and (28), the proof of the theorem is completed. □

If $r = 1$, we have that $g_r^{(x)}(\cdot;k_1)$ is independent of $x \in S_m$ (see Corollary 4). By induction, assume that $g_r^{(x)}(\cdot)$ is independent of $x \in S_m$. Clearly, from the above relation, we have that $g_{r+1}^{(x)}(\cdot;k_1)$ can be expressed as weighted sums of $g_i^{(x)}(\cdot;k_1)$ for $i = 1, 2, \dots, r$. Since the right hand side of the above relation does not involve any $x \in S_m$, $g_{r+1}^{(x)}(\cdot;k_1)$ must be independent of x. Therefore, from now on we will drop the superscript x from $g_r^{(x)}(\cdot;k_1)$. Hence we have the following corollary.

Corollary 6: For any $x \in S_m$, the probability $g_r^{(x)}(n;k_1) = \mathbb{P}_x(N_r(k_1) = n)$ is independent of x and satisfies the recurrence relation

$$\begin{aligned}
&g_r(n;k_1) \\
&= \sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j g_r(n-j-1;k_1) + \Big(p_{2^m-1}\Big)^{k_2+(r-1)(l+1)} 1_n(k_2+(r-1)(l+1)+1) \\
&\quad + \sum_{j_1=0}^{r-2}\sum_{j_2=0}^{l} q_{2^m-1}\Big(p_{2^m-1}\Big)^{k_2+j_1(l+1)+j_2} g_{r-1-j_1}(n-1-k_2-j_1(l+1)-j_2;k_1). \qquad (29)
\end{aligned}$$

We now derive the generating function $\zeta_r(s;k_1)$ of $N_r(k_1)$ using the recurrence relation. For $r = 1$, we have already obtained the expression of $\zeta_1(s;k_1)$ (see Corollary 5). For $r \geq 2$, we can't directly obtain the expression of $\zeta_r(s;k_1)$. Instead, we will obtain a recurrence relation in terms of the generating functions. Indeed, for $r \geq 2$, we have

$$\begin{aligned}
\zeta_r(s;k_1) &= s\Big(p_{2^m-1}s\Big)^{k_2+(r-1)(l+1)} + \sum_{n=0}^{\infty}\sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j g_r(n-1-j;k_1)s^n \\
&\quad + \sum_{n=0}^{\infty}\sum_{j_1=0}^{r-2}\sum_{j_2=0}^{l} q_{2^m-1}\Big(p_{2^m-1}\Big)^{k_2+j_1(l+1)+j_2} \\
&\qquad \times g_{r-1-j_1}(n-1-k_2-j_1(l+1)-j_2;k_1)s^n \\
&= s\Big(p_{2^m-1}s\Big)^{k_2+(r-1)(l+1)} + \sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}\zeta_r(s;k_1) \\
&\quad + \sum_{j_1=0}^{r-2}\sum_{j_2=0}^{l} q_{2^m-1}s\Big(p_{2^m-1}s\Big)^{k_2+j_1(l+1)+j_2}\zeta_{r-1-j_1}(s;k_1). \qquad (30)
\end{aligned}$$

Simplifying equation (30), we obtain a recurrence relation involving $\zeta_r(s;k_1)$. This is given in the following lemma.

Lemma 1: For $r \geq 2$, the sequence of the probability generating functions satisfies the following recurrence relation

$$\begin{aligned}
&\Big(1 - \sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}\Big)\zeta_r(s;k_1) \\
&= \Big(p_{2^m-1}s\Big)^{k_2}\Big(\sum_{j=0}^{l} q_{2^m-1}s\Big(p_{2^m-1}s\Big)^j\Big)\sum_{j_1=0}^{r-2}\Big(p_{2^m-1}s\Big)^{j_1(l+1)}\zeta_{r-1-j_1}(s;k_1) \\
&\quad + s\Big(p_{2^m-1}s\Big)^{k_2+(r-1)(l+1)}.
\end{aligned} \tag{31}$$

Now, we are ready to prove the main theorem, namely Theorem 1.

Proof: The generating function of the sequence $\{\zeta_r(s;k_1) : r \geq 1\}$ is denoted by $\Xi(z;k_1)$, *i.e.*,

$$\Xi(z;k_1) = \sum_{r=1}^{\infty} \zeta_r(s;k_1) z^r.$$

Now, using (31) we obtain the generating function $\Xi(z;k_1)$ as follows:

$$\begin{aligned}
&\Big(1 - \sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}\Big)\Xi(z;k_1) \\
&= \sum_{r=1}^{\infty}\Big(1 - \sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}\Big)\zeta_r(s;k_1)z^r \\
&= \Big(1 - \sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}\Big)\zeta_1(s;k_1)z + \sum_{r=2}^{\infty} s\Big(p_{2^m-1}s\Big)^{k_2+(r-1)(l+1)} z^r \\
&\quad + \sum_{r=2}^{\infty}\Big(p_{2^m-1}s\Big)^{k_2}\Big(\sum_{j=0}^{l} q_{2^m-1}s\Big(p_{2^m-1}s\Big)^j\Big)\sum_{j_1=0}^{r-2}\Big(p_{2^m-1}s\Big)^{j_1(l+1)}\zeta_{r-1-j_1}(s;k_1)z^r \\
&= sz\Big(p_{2^m-1}s\Big)^{k_2} + sz\Big(p_{2^m-1}s\Big)^{k_2}\sum_{r=1}^{\infty}\Big(p_{2^m-1}s\Big)^{r(l+1)} z^r \\
&\quad + \Big(p_{2^m-1}s\Big)^{k_2}\Big(\sum_{j=0}^{l} q_{2^m-1}s\Big(p_{2^m-1}s\Big)^j\Big)\sum_{j_1=0}^{\infty}\Big(p_{2^m-1}s\Big)^{(l+1)j_1}\sum_{r=j_1}^{\infty}\zeta_{r-j_1+1}(s;k_1)z^{r+2} \\
&= \frac{sz\Big(p_{2^m-1}s\Big)^{k_2}}{1-\Big(p_{2^m-1}s\Big)^{(l+1)}z} + \frac{\Big(p_{2^m-1}s\Big)^{k_2}\Big(\sum_{j=0}^{l} q_{2^m-1}s\Big(p_{2^m-1}s\Big)^j\Big)z\Xi(z;k_1)}{1-\Big(p_{2^m-1}s\Big)^{(l+1)}z}.
\end{aligned} \tag{32}$$

Now, from the above equation (32), we can easily solve $\Xi(z;k_1)$ to obtain

$$\begin{aligned}
\Xi(z;k_1) &= \left[sz\Big(p_{2^m-1}s\Big)^{k_2}\right]\Bigg[\Big(1-\sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}\Big) \\
&\quad \times \Big(1-\Big(p_{2^m-1}s\Big)^{(l+1)}z\Big) - z\Big(p_{2^m-1}s\Big)^{k_2}\Big(\sum_{j=0}^{l} q_{2^m-1}s\Big(p_{2^m-1}s\Big)^j\Big)\Bigg]^{-1} \\
&= \frac{zs\Big(p_{2^m-1}s\Big)^{k_2}}{1-\sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}}\Bigg[1-\Big(p_{2^m-1}s\Big)^{(l+1)}z \\
&\quad - \frac{z\Big(p_{2^m-1}s\Big)^{k_2}\Big(\sum_{j=0}^{l} q_{2^m-1}s\Big(p_{2^m-1}s\Big)^j\Big)}{1-\sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}}\Bigg]^{-1} \\
&= \frac{zs\Big(p_{2^m-1}s\Big)^{k_2}}{1-\sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}} \\
&\quad \times \left[1-z\left(\Big(p_{2^m-1}s\Big)^{(l+1)} + \frac{\Big(p_{2^m-1}s\Big)^{k_2}\Big(\sum_{j=0}^{l} q_{2^m-1}s\Big(p_{2^m-1}s\Big)^j\Big)}{1-\sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}}\right)\right]^{-1}. \qquad (33)
\end{aligned}$$

Now, we obtain $\zeta_r(s;k_1)$ by calculating the coefficient of z^r in the equation (33). Observe that coefficient of z^r is obtained by multiplying the coefficient of z^{r-1} in the expression in the last line in (33) by $s\Big(p_{2^m-1}s\Big)^{k_2}/\Big(1-\sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}\Big)$. Using the expansion $(1-az)^{-1}=\sum_{t=0}^{\infty} a^t z^t$, we have

$$\begin{aligned}
\zeta_r(s;k_1) &= \frac{s\Big(p_{2^m-1}s\Big)^{k_2}}{1-\sum_{j=0}^{k_2-1} q_{2^m-1}s\Big(p_{2^m-1}s\Big)^j}\Bigg[\Big(p_{2^m-1}s\Big)^{l+1} \\
&\quad + \frac{s\Big(p_{2^m-1}s\Big)^{k_2}}{1-\sum_{j=0}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^j s^{j+1}}\sum_{j=0}^{l} q_{2^m-1}\Big(p_{2^m-1}s\Big)^j\Bigg]^{r-1}.
\end{aligned}$$

This completes the proof of theorem 1. □

6. Conclusion

In this article we have defined a new discrete distribution, called *generalized Binomial type distribution*. The probability generating function of the distribution has been given along with its connections to the classical Binomial distribution as well as the negative Binomial distribution. We have studied the number of overlapping runs of length k_1 until the r^{th} occurrence of l-look-back run of length k ($k_1 \leq k$) for the m^{th} order Markov chain and obtained the explicit expression of its probability generating function. Further, we have shown that our result generalizes the results of Aki and Hirano (1994) when we consider $r=1$ for both *i.i.d.* as well as Markov dependent case. Since our stopping time is quite general, our theorem will also provide similar results when we apply it to different cases such as the r^{th} occurrence of non-overlapping runs or r^{th} occurrence of overlapping runs or r^{th} occurrence of μ-overlapping runs (for positive μ).

Our result shows that the conditional distribution, that we have considered, has a renewal structure (see Feller (1968)) in the sense that it splits into independent sums of random variables, which may be interpreted as arrival times in a renewal process. Further, it is also seen that the arrival times are identical except the first arrival time. In other words, it admits a delayed renewal structure. We are able to identify the arrival times through the newly defined generalized Binomial type distribution and extended geometric discrete distribution. This renewal structure, in turn, can also be used to obtain approximate distribution of number of runs when the value of r is large. For instance, we may obtain the strong law of large numbers for the number of overlapping runs of length k_1.

We also provide a versatile method of proving the result where we convert our problem from the m^{th} order Markov chain into a simple Markov chain by combining the states. This allows us to use the Markov chain machinery, namely the strong Markov property, to derive the recurrence relation and use the method of generating functions effectively to obtain our results. Our method is quite powerful and can be used to prove similar results for other run statistic. We expect that, in future, there will be more applications of our method.

Acknowledgement

The author wishes to thank the referees for their helpful comments which have improved the presentation of the paper.

References

Aki, S. (1985). Discrete distributions of order k on a binary sequence. *Annals of the Institute of Statistical Mathematics*, **37**, 205–224.

Aki, S. and Hirano, K. (1994). Distributions of numbers of failures and successes until the first consecutive k successes. *Annals of the Institute of Statistical Mathematics*, **46**, 193–202.

Aki, S. and Hirano, K. (1995). Joint distributions of numbers of success-runs and failures until the first k consecutive successes. *Annals of the Institute of Statistical Mathematics*, **47**, 225–235.

Aki, S. and Hirano, K. (2000). Number of success-runs of specified length until certain stopping time rules and generalized binomial distributions of order k. *Annals of the Institute of Statistical Mathematics*, **52**, 767–777.

Anuradha. (2023). Asymptotic results for generalized runs in higher order Markov Chains. *Statistics and Applications*, **21**, 189–207.

Chadjiconstantinidis, S. and Koutras, M. V. (2001). Distributions of the numbers of failures and successes in a waiting time problem. *Annals of the Institute of Statistical Mathematics*, **53**, 576–598.

Ebneshahrashoob, M. and Sobel, M. (1990). Sooner or Later waiting time problems for Bernoulli trials: Frequency and run quotas. *Statistics and Probability Letters*, **9**, 5–11.

Feller, W. (1968). *An Introduction to Probability Theory and its Applications. Vol - I.* John Wiley, New York, 3rd ed.

Fu, J. C. and Koutras, M. V. (1994). Distribution Theory of runs: a Markov Chain approach. *Journal of the American Statistical Association*, **89**, 1050–1058.

Han, Q. and Aki, S. (2000). A unified approach to binomial type distributions of order k. *Communications in Statistics - Theory and Methods*, **29**, 1929–1943.

Hirano, K., Aki, S., and Uchida, M. (1997). Distributions of numbers of success-runs until the first consecutive k successes in higher order Markov dependent trials. In: Balakrishnan, N. (eds) *Advances in Combinatorial Methods and Applications to Probability and Statistics*, Birkhäuser Boston, Boston, 401–410.

Inoue, K. and Aki, S. (2002). Generalized binomial and negative binomial distributions of order k by the l-overlapping enumeration scheme. *Annals of the Institute of Statistical Mathematics*, **55**, 153–167.

Koutras, M. V. (1997). Waiting time distributions associated with runs of fixed length in two-state Markov chains. *Annals of the Institute of Statistical Mathematics*, **49**, 123–139.

Makri, F. S. and Philippou, A. N. (2005). On binomial and circular binomial distributions of order k for l- overlapping success runs of length k. *Statistical Papers*, **46**, 411–432.

Makri, F. S., Philippou, A. N., and Psillakis, Z. M. (2007). Polya, inverse Polya, and circular Polya distributions of order k for l- overlapping success runs. *Communications in Statistics - Theory and Methods*, **36**, 657–668.

Philippou, A. N. and Makri, F. S. (1986). Successes, runs and longest runs. *Statistics and Probability Letters*, **4**, 211–215.

Uchida, M. (1998). On number of occurrences of success runs of specified length in a higher order two-state Markov chain. *Annals of the Institute of Statistical Mathematics*, **50**, 587–601.