Statistics and Applications {ISSN 2454-7395 (online)}
Volume 22, No. 1, 2024 (New Series), pp 83–100
http://www.ssca.org.in/journal

Distribution of Runs of Length Exactly k_1 Until a Stopping Time for Higher Order Markov Chain

Anuradha
Department of Statistics, Lady Shri Ram College for Women, University of Delhi.

Received: 02 November 2022; Revised: 05 June 2023; Accepted: 10 June 2023

Abstract

Consider an m^{th} order Markov chain $\{X_j : j \geq -m+1\}$ taking values in $\{\mathbf{0}, \mathbf{1}\}$. We set $R_i = 0$ for $i = 0, -1, \ldots, -l+1$. A l-look-back run of length k starting at i, R_i is defined inductively as a run of $\mathbf{1}$'s starting at i, provided that no l-look-back run of length k occurs, starting at time $i-1, i-2, \ldots, i-l$, *i.e.*, $R_i = \prod_{j=i-1}^{i-l}(1-R_j)\prod_{j=i}^{i+k-1} X_j$. We study the conditional distribution of the number of runs of length exactly k_1, till the r-th occurrence of the l-look-back run of length k where $k_1 \leq k-1$ and obtain the explicit expression of it's probability generating function. We establish that the number of runs can be written as sum of r independent random variables with the first term having a slightly different distribution. We further establish the strong law of large numbers for the number of runs of length exactly k_1.

Key words: Runs; Markov chain; Stopping time; Probability generating function; Strong Markov property: Strong law of large numbers.

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

1. Introduction

Theory of distributions of runs has been studied, since Feller (1968) introduced runs as an example of a renewal event. In recent years, this field has received a lot of interest among researchers. Many powerful techniques such as Markov embedding technique, method of conditional p.g.f.s *etc.* have been developed which enabled us to study new features of the distributions of various run statistics. For a more detailed discussion on the run statistics and its application, we refer the readers to Balakrishnan and Koutras [2002].

We consider an m-th order homogeneous $\{\mathbf{0}, \mathbf{1}\}$-valued Markov chain. Further, we assume that the initial condition $\{X_0 = x_0, X_{-1} = x_1, \ldots, X_{-m+1} = x_{m-1}\}$ is given to us. The state $\mathbf{1}$ can be thought as success in an experiment while $\mathbf{0}$ as failure. A run of length k is a consecutive occurrence of k successes. Anuradha (2022) introduced the l-look-back counting scheme for runs. In this scheme a run is counted starting at time i, if $X_i = X_{i+1} = \cdots = X_{i+k-1} = \mathbf{1}$, and no runs can be counted till the time point $i+l$.

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

The next counting of run can start only from the time point $i + l + 1$. This mechanism is repeated every time a run is counted. In other words, if a run is counted starting at time i, there are k-consecutive successes from the time point i and no runs of length k has been counted which had the starting time as $i-1, i-2, \ldots, i-l$. Such a run will be referred as a l-look-back run of length k. Clearly, if $l = 0$, this counting of run matches exactly with the number of overlapping runs of length k, while if we set $l = k-1$, this counting results in the number of non-overlapping runs of length k. Aki and Hirano (2000) also defined a counting scheme which they referred as μ-overlapping counting. It should be noted that both these concepts match if we set $l = k - \mu - 1$.

The following example illustrates the practical usage of the l-look-back counting scheme for runs of length k. Consider an experiment of a drug administration where observations are taken every hour for the presence or absence (success or failure) of a particular symptom, say, fever exceeding a specified temperature. If we observe the presence of the symptom for k-successive time points, a drug has to be administered; however, as is the case with most drugs, once the drug is administered, we have to wait for l-hours for the next administration of the drug. But the process of the observation for the presence or absence of the symptom is continued as ever. In such a case, the number of administrations of the drug until time point n, is the number of l-look-back runs of length k up to time n.

Aki and Hirano (1994) studied the marginal distributions of failures, successes and success-runs of length less than k until the first occurrence of consecutive k successes where the underlying random variables are either *i.i.d.* or homogeneous Markov chain or binary sequence of order k. Aki and Hirano (1995) derived the joint distributions of failures, successes and success-runs for the same set-up. Hirano *et. al.* (1997) obtained the distributions of number of success-runs of a specific length for various counting schemes (*e.g.* runs of length k_1, overlapping runs of length k_1, non-overlapping runs of length k_1 *etc.*) until the first occurrence of the success-run of length k for a m-th order homogeneous Markov chain where $m \leq k_1 < k$. Uchida (1998) studied the joint distributions of the waiting time and the number of outcomes such as failures, successes and success-runs of length less than k for various counting schemes of runs for an m^{th} order homogeneous Markov chain. Chadjiconstantindis and Koutras (2001) also obtained the distribution of number of failures and successes in a waiting time problem.

In this paper, we study the distribution of runs of successes of exact length. A run of length exactly k can be described as an occurrence of a failure, followed by k consecutive successes, followed by another failure. The literature on runs of exact length is rather limited. This is indeed a difficult problem, specially when the underlying distribution of random variables has a dependent structure.

In recent years, the runs of exact length has found usage in very important areas. We site one such example here. The study of random sequences constitutes an important part of cryptography specially in the areas of challenge and response authentication systems, generation of digital signatures, and zero-knowledge protocols. Many protocols in cryptography depend on the assumption that the resulting ciphertext from a cipher (cryptographic algorithm) should appear to be as random as possible. Various tests are used for testing the randomness of such ciphertexts, which in turn help in deciding whether a given protocol leaks information or not. Doganaksoy *et. al.* (2015) developed three statistical randomness tests based on runs of exact length and named them as runs of length one, runs of length two, and runs of length three tests respectively and showed that they work better than the traditional tests. However, the main challenge in the wider application of their work was that the distribution of the resulting statistic is not tractable when the (exact) length of run is large. In fact they could use only lengths 1, 2 and 3. Hence there is an imperative need to study the distribution, or at least find good approximation of the distribution, of runs of exact length for larger values of length, specially when the underlying random variables are not *i.i.d.* but have some dependent structure.

Since the study of the exact distribution of runs of exact length is complicated, it is prudent to find a simpler structure embedded in this set up. We study the conditional distribution of the number of runs of length exactly k_1, until a specified stopping time, namely the r^{th} occurrence of the l-look-back run of length k where $k_1 < k$. The study of distributions of runs until a stopping time brings out many salient features of various run statistic and establishes new connection between various discrete distributions. Indeed, our results exhibit an independence structure in the number of runs until the stopping time where we may explicitly write the distribution in terms of simpler random variables following Bernoulli and geometric distributions. (see Corollary 1 for details).

The novelty of our method lies in translating our problem into a first order homogeneous Markov chain. Indeed, we define a new first order Markov chain taking values in a finite set in such a way that the states of the new chain combines the last k_1 states of the previous chain (refer to the third section for exact definition). Further, the states of the original m-th order Markov chain may be recovered from the states of the newly defined Markov chain. This allows us to translate the problem in terms of the new Markov chain. For a simple Markov chain, the powerful results such as the strong Markov property can now be used to derive a recurrence relation between the probabilities. We now employ the method of conditional probability generating functions. We use this basic relation involving the probabilities to obtain a recurrence relation involving probability generating functions. This, in turn, provides a simple linear equation involving the generating function of the probability generating functions which can be solved to obtain its expression.

The explicit expression of the probability generating function implies that the distribution of the number of runs of length exactly k_1 until the stopping time has a renewal structure. Hence the number of runs until the stopping time splits into sum of independent random variables, which may be interpreted as arrival times in a renewal process. Further, we have shown that the arrival times are identical except the first arrival time. In other words, it admits a delayed renewal structure. We are also able to identify the arrival times through geometric and Bernoulli random variables. Thus we are able to approximate the number of runs of length exactly k_1 through simpler random variables.

We may apply our results to obtain an approximation of the number of runs of length exactly k_1 until time n in the following way: we choose some $k > k_1$ and find the number of non-overlapping runs of length k, *i.e.*, number of $(k-1)$-look-back runs of length k, until time n, say r. Clearly the total number of runs of length exactly k_1 until time n lies between the number of runs of length exactly k_1 until the r^{th} and $(r+1)^{th}$ occurrence of non-overlapping runs of length k respectively. Now we use our main result to compute the distribution of the number of runs of length exactly k_1 until this r^{th} as well as $(r+1)^{th}$ occurrences of non-overlapping runs of length k. For large values of n, this works quite well. We use the Markov inequality and this method of approximation effectively to derive the strong law of large numbers (see Theorem 2) for the number of runs of length exactly k_1. We hope that the methods of approximation can, in future, be extended to obtain a central limit theorem as well as the law of iterated logarithm for the same.

In the next section, we give the important definitions and state the main Theorem and Corollary related to the distribution of the number of runs of length exactly k_1 until a stopping time, where $k_1 < k$. Section 3 is devoted towards formalizing the underlying set up for deriving the results. In Section 4, we prove the main theorem, while in Section 5, we prove the strong law for the number of runs.

2. Definitions and statement of results

Let $X_{-m+1}, X_{-m+2}, \ldots, X_0, X_1, \ldots$ be a sequence of stationary m-order $\{\mathbf{0}, \mathbf{1}\}$ valued Markov chain. It is assumed that the states of $X_{-m+1}, X_{-m+2}, \ldots, X_0$ are known, *i.e.*, we are given the initial condition $\{X_0 = x_0, X_{-1} = x_1, \ldots, X_{-m+1} = x_{m-1}\}$.

To make things formal, for any $i \geq 0$, define $C_i = \{0, 1, \ldots, 2^i - 1\}$. It is clear that C_i and $\{\mathbf{0}, \mathbf{1}\}^i$ can be identified easily by the 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 m^{th} order Markov chain, we have, for any $n \geq 0$,

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

where $x = \sum_{j=0}^{m-1} 2^j x_j \in C_m$. Consequently, 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 $x \in C_i$. One particular case will be of importance in our study, when $\{X_n = 1, X_{n-1} = 1, \ldots, X_{n-m+1} = 1\}$. In our notation, this condition will become $x = \sum_{j=0}^{m-1} 2^j 1 = 2^m - 1$. Thus, using (1), for all $n \geq 0$, we have

$$p_{2^m-1} = \mathbb{P}(X_{n+1} = 1 | X_n = 1, X_{n-1} = 1, \ldots, X_{n-m+1} = 1) = 1 - q_{2^m-1}.$$

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

$$R_i(k,l) = \prod_{j=i-l}^{i-1} (1 - R_j(k,l)) \prod_{j=i}^{i+k-1} X_j. \tag{2}$$

If $R_i(k,l) = 1$, we say that an *l-look-back run of length k* has been recorded which started at time i.

It should be noted that for an 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 an l-look-back run of length k.

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

Definition 2: For $r \geq 1$, the stopping time $\tau_r(k,l)$ be the (random) time point at which the r-th occurrence of l-look-back run of length k is completed. In other words,

$$\tau_r(k,l) = k - 1 + \inf\{n : \sum_{i=1}^{n} R_i(k,l) = r\}. \tag{3}$$

Now we define the runs of length exactly k.

Definition 3: When $k(\geq 1)$ consecutive successes, either occur at the beginning of the sequence or end of the sequence or bordered on both sides by failures, contribute towards the counting of a run then we call it run of length exactly k.Note that when there are more than k consecutive successes then it is not counted as run of length exactly k.

We may represent this mathematically as follows:

$$\epsilon_i(k) = \begin{cases} \prod_{j=1}^{k} X_j(1 - X_{k+1}) & \text{if } i = 1 \\ (1 - X_{i-1}) \prod_{j=i}^{i+k-1} X_j(1 - X_{i+k}) & \text{if } 1 < i < n - k + 1 \\ (1 - X_{n-k}) \prod_{j=n-k+1}^{n} X_j & \text{if } i = n - k + 1. \end{cases}$$

Note here that $\epsilon_i(k) = 1$ if and only if a run of length exactly k starts at time point i. Now, we define the total number of runs of length exactly k by

$$N_n(k) = \sum_{i=1}^{n} \epsilon_i(k). \tag{4}$$

In this paper, we study the number of runs of length exactly k till the stopping time $\tau_r(k,l)$ (see Definition (2)). Fix any constant $k_1 \leq k$. For each $r \geq 1$, we define the random variable

$$N_r(= N_r(k_1)) := N_{\tau_r(k,l)}(k_1) = \sum_{i=1}^{\tau_r(k,l)} R_i(k_1) \tag{5}$$

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

Before we proceed, we present an example to facilitate the understanding. Consider the following sequence of **0**'s and **1**'s of length 20

11010111011111011101.

For $k = 3$ and $l = 1$, it should be noted that, $R_6(3,1) = R_{10}(3,1) = R_{12}(3,1) = R_{16}(3,1) = 1$, while for other values of i, $R_i(3,1) = 0$. Thus, $\tau_1(3,1) = 8, \tau_2(3,1) = 12, \tau_3(3,1) = 14$ and $\tau_4(3,1) = 18$. For $k_1 = 2$, the number of runs of length exactly k_1 are given by $N_1 = N_2 = N_3 = N_4 = 1$ respectively.

Let us define the probability generating function of N_r, *i.e.*,

$$\zeta_r(s;k_1) := \sum_{n=0}^{\infty} \mathbb{P}(N_r = n)s^n. \tag{6}$$

Theorem 1: For any initial condition $x \in C_i$, $k_2 = k - k_1 > 0$ and $k_1 \geq m$, the probability generating function of N_r is given by

$$\zeta_r(s; k_1) = \left[\frac{\left(p_{2^m-1}\right)^{k_2}}{q_{2^m-1} + \left(p_{2^m-1}\right)^{k_2} - q_{2^m-1}s}\right]\left[\left(p_{2^m-1}\right)^{l+1} + \frac{\left(p_{2^m-1}\right)^{k_2}}{q_{2^m-1} + \left(p_{2^m-1}\right)^{k_2} - q_{2^m-1}s}\left(1 - \left(p_{2^m-1}\right)^{l+1}\right)\right]^{r-1}.$$

Theorem 1 provides a useful representation of N_r in terms of Bernoulli and geometric random variables when $k_2 > 0$. Let us set,

$$p_E = \frac{\left(p_{2^m-1}\right)^{k_2}}{1 - \sum_{t=1}^{k_2-1}\left(p_{2^m-1}\right)^t q_{2^m-1}} = \frac{\left(p_{2^m-1}\right)^{k_2}}{q_{2^m-1} + \left(p_{2^m-1}\right)^{k_2}}. \tag{7}$$

Corollary 1: Let $\{G_i : i = 1, \ldots, r\}$ and $\{B_i : i = 1, \ldots, r\}$ be two independent sets of random variables with each G_i having a geometric distribution (taking values in $\{0, 1, \ldots, \}$) with parameter p_E and each B_i having a Bernoulli distribution with parameter $\left(1 - \left(p_{2^m-1}\right)^{l+1}\right)$, then we have

$$N_r \stackrel{d}{=} G_1 + \sum_{i=2}^{r} G_i B_i. \tag{8}$$

Indeed, it is easy to see that the probability generating function of G_i, for $i \geq 1$, is given by

$$\frac{\left(p_{2^m-1}\right)^{k_2}}{q_{2^m-1} + \left(p_{2^m-1}\right)^{k_2} - q_{2^m-1}s}$$

and the probability generating function of B_i, for $i \geq 1$, is given by

$$\left(p_{2^m-1}\right)^{l+1} + s\left(1 - \left(p_{2^m-1}\right)^{l+1}\right).$$

Therefore, the probability generating function of G_iB_i is given by

$$\left(p_{2^m-1}\right)^{l+1} + \frac{\left(p_{2^m-1}\right)^{k_2}}{q_{2^m-1} + \left(p_{2^m-1}\right)^{k_2} - q_{2^m-1}s}\left(1 - \left(p_{2^m-1}\right)^{l+1}\right). \tag{9}$$

From the independence of G_i and B_i for $i \geq 1$, the corollary easily follows.

When $k_2 = 0$, *i.e.*, $k = k_1$, we can obtain the the probability generating function, but it is difficult to identify the exact distribution (see Section 4).

The delayed renewal structure of the number of runs of exact length until the stopping time, observed in equation (8), can be used for approximating the original distribution when the number of trials are large. Indeed we may obtain a strong law for the number of runs of exact length using this.

Let us set $k = k_1 + 1$ and $l = k - 1 = k_1$. Then, the expectation of G_1B_1 can be easily computed from the expression of the probability generating function in (9). Indeed, it is given by

$$\mu_1 = \frac{q_{2^m-1}}{p_{2^m-1}}\Big(1 - \big(p_{2^m-1}\big)^{k_1+1}\Big). \tag{10}$$

We will further define a constant μ. Let S be the first time when k successive heads have occurred given the initial condition of k successive heads. In section 5, We will show that S is finite with probability 1. Further, its expectation is also finite. We denote

$$\mu = \mathbb{E}(S). \tag{11}$$

Theorem 2: For any initial condition $x \in C_i$ and $k_1 \geq m$, we have

$$\frac{1}{n}N_n(k_1) \to \frac{\mu_1}{\mu}$$

as $n \to \infty$ with probability 1.

3. Formal set-up

In this section, we outline the basic set up which will be used in the subsequent section to establish the results. Let us define two functions $f_0, f_1 : C_{k_1} \to C_{k_1}$ by

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

Further define a projection $\theta_m : C_{k_1} \to C_m$ 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 C_{k_1}. Further, the random variables X_n's are stationary and forms a m^{th} order Markov chain, hence we have that $\{Y_n : n \geq 0\}$ is a homogeneous Markov chain with transition matrix given by

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

It should be noted that Y_n is even if and only if $X_n = 0$. This motivates us to define the function $\kappa : C_{k_1} \to \{0, 1\}$ 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 $x \in C_m$. We denote the probability measure governing the distribution of $\{Y_n : n \geq 1\}$ with $Y_0 = x \in C_k$ 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

$$T := \inf\{i \geq k_1 : \prod_{j=i-k_1+1}^{i} X_j = 1\}. \tag{12}$$

We would like to translate the above definition in terms of 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 and this has not happened earlier. So, T can be better described as

$$T = \inf\{i \geq k_1 : Y_i = 2^{k_1} - 1\}, \tag{13}$$

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 $u \in C_m$, 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 happens, then k_1 consecutive successes must have occured previously since $k_1 \leq k$. Therefore, we have $P_x(T < \tau_1(k,l)) = 1$.

4. Number of runs of exact length until stopping time

First we establish the basic recurrence relation which is central to our result. Define the probability $g_r^{(x)}(n)$ by

$$g_r^{(x)}(n) := \mathbb{P}_x(N_r = n) \tag{14}$$

for $n \in \mathbb{Z}$. We note that since $N_r \geq 0$, $\mathbb{P}_x(N_r = n) = 0$ for $n < 0$. Our first task is to show that $g_r^{(x)}(n)$ is independent of x.

Theorem 3: Suppose that $k_2 = k - k_1 > 0$. For any $x \in C_{k_1}$ and any $n \geq 0$, we have

$$\begin{aligned} g_1^{(x)}(n) = q_{2^m-1} g_1^{(2^m-2)}(n-1) &+ \sum_{t=1}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^t g_1^{(2^m-2)}(n) \\ &+ \Big(p_{2^m-1}\Big)^{k_2} \mathbb{I}_n(0) \end{aligned} \tag{15}$$

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

$$\mathbb{I}_{u_1}(u_2) = \begin{cases} 1 & \text{if } u_1 = u_2 \\ 0 & \text{otherwise.} \end{cases}$$

Proof: When $k_2 = k - k_1 > 0$ and $r = 1$, we have

$$\begin{aligned} g_1^{(x)}(n) = \mathbb{P}_x(N_1 = n) &= \mathbb{P}_x(N_1 = n, Y_{T+1} = 2^{k_1} - 2) \\ &+ \sum_{t=1}^{k_2-1} \mathbb{P}_x(N_1 = n, Y_{T+1} = 2^{k_1} - 1, \ldots, Y_{T+t} = 2^{k_1} - 1, Y_{T+t+1} = 2^{k_1} - 2) \\ &+ \mathbb{P}_x(N_1 = n, Y_{T+1} = 2^{k_1} - 1, Y_{T+2} = 2^{k_1} - 1, \ldots, \\ &\qquad Y_{T+k_2-1} = 2^{k_1} - 1, Y_{T+k_2} = 2^{k_1} - 1). \end{aligned} \tag{16}$$

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

$$\begin{aligned} &\mathbb{P}_x(N_1 = n, Y_{T+1} = 2^{k_1} - 1, Y_{T+2} = 2^{k_1} - 1, \ldots, \\ &\qquad Y_{T+t} = 2^{k_1} - 1, Y_{T+t+1} = 2^{k_1} - 2) \\ &= \mathbb{P}_x(N_1 = n \mid Y_{T+1} = 2^{k_1} - 1, Y_{T+2} = 2^{k_1} - 1, \ldots, \\ &\qquad Y_{T+t} = 2^{k_1} - 1, Y_{T+t+1} = 2^{k_1} - 2) \\ &\times \mathbb{P}_x(Y_{T+1} = 2^{k_1} - 1, Y_{T+2} = 2^{k_1} - 1, \ldots, \\ &\qquad Y_{T+t} = 2^{k_1} - 1, Y_{T+t+1} = 2^{k_1} - 2). \end{aligned} \tag{17}$$

The second term in (17) can be written as

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

Now, for any $1 \leq j \leq t, T + j - 1$ is also a stopping time. 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, $\{Y_{T+1} = 2^{k_1} - 1, \ldots, Y_{T+j-1} = 2^{k_1} - 1\} \in \mathcal{F}_{T+j-1}$ and $\{Y_{T+j} = 2^{k_1} - 1\} \in \mathcal{F}_{(T+j-1)+}$. Thus, using the strong Markov property, we can write

$$\begin{aligned} &\mathbb{P}_x(Y_{T+j} = 2^{k_1} - 1 \mid Y_{T+1} = 2^{k_1} - 1, \ldots, Y_{T+j-1} = 2^{k_1} - 1) \\ &= \mathbb{P}_{Y_{T+j-1}}(Y_{T+j} = 2^{k_1} - 1) = \mathbb{P}_{2^{k_1}-1}(Y_1 = 2^{k_1} - 1) = p_{2^m-1}. \end{aligned} \tag{18}$$

A similar argument shows that

$$\mathbb{P}_x(Y_{T+t+1} = 2^{k_1} - 2 \mid Y_{T+1} = 2^{k_1} - 1, \ldots, Y_{T+t} = 2^{k_1} - 1) = q_{2^m-1}. \tag{19}$$

For the first term in (17), we note that $T + t + 1$ is also a stopping time and $\{Y_{T+1} = 2^{k_1} - 1, \ldots, Y_{T+t} = 2^{k_1} - 1, Y_{T+t+1} = 2^{k_1} - 2\} \in \mathcal{F}_{T+t+1}$. Since $Y_T = 2^{k_1} - 1$, we must have either $X_{T-k_1} = 0$ and $X_{T-j} = 1$ for $j = 0, 1, \ldots, k_1 - 1$ or $T = k_1$. Further, since $Y_{T+j} = 2^{k_1} - 1$ for $j = 1, \ldots, t$ and $Y_{T+t+1} = 2^{k_1} - 2$, we also have $X_{T+j} = 1$ for $j = 0, 1, \ldots, t$ 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 0 runs of length exactly 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 runs of length exactly k_1 up to time $T+t+1$ is 0. Since $t \le k_2 - 1$, we have that $T+t+1 < \tau_1(k,l)$. Let us define $Y'_i = Y_{i+T+t+1}$ for $i \ge 0$. Now, using the strong Markov property, we have that $\{Y'_i : i \ge 0\}$ is a homogeneous Markov chain with same transition matrix as that of $\{Y_i : i \ge 0\}$ with $Y'_0 = 2^{k_1} - 2$. Now, define $\tau'_1(k,l)$ as the stopping time for the process $\{Y'_i : i \ge 0\}$. From the above discussion, we have that $\tau_1(k,l) = T+t+1+\tau'_1(k,l)$. Further, if we define, N'_1 as the number runs of length exactly k_1 up to time $\tau'_1(k,l)$ for the process $\{Y'_i : i \ge 0\}$, we must have that $N'_1 = n$. Therefore, we have,

$$\begin{aligned}
&\mathbb{P}_x(N_1 = n \mid Y_{T+1} = 2^{k_1}-1, \ldots, Y_{T+t} = 2^{k_1}-1, Y_{T+t+1} = 2^{k_1}-2) \\
&= \mathbb{P}_{(2^m-2)}(N'_1 = n) = g_1^{(2^m-2)}(n). \qquad (20)
\end{aligned}$$

Now, the first term in (16) can be written as

$$\begin{aligned}
&\mathbb{P}_x(N_1 = n, Y_{T+1} = 2^{k_1}-2) \\
&= \mathbb{P}_x(N_1 = n \mid Y_{T+1} = 2^{k_1}-2, Y_T = 2^{k_1}-1)\mathbb{P}_x(Y_{T+1} = 2^{k_1}-2 \mid Y_T = 2^{k_1}-1) \\
&= q_{2^m-1}\mathbb{P}_x(N_1 = n \mid Y_{T+1} = 2^{k_1}-2, Y_T = 2^{k_1}-1). \qquad (21)
\end{aligned}$$

The arguments leading to equation (20) can now be repeated to conclude that

$$\mathbb{P}_x(N_1 = n \mid Y_{T+1} = 2^{k_1}-2, Y_T = 2^{k_1}-1) = \mathbb{P}_{(2^m-2)}(N_1 = n-1) = g_1^{(2^m-2)}(n-1). \qquad (22)$$

Using the equivalent characterisation of T (see equation (13)) we note that $Y_T = 2^{k_1} - 1$ with probability 1. Hence, for the last term in (16) becomes

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

Note that given $\{Y_{T+1} = 2^{k_1}-1, \ldots, Y_{T+k_2-1} = 2^{k_1}-1, Y_{T+k_2} = 2^{k_1}-1\}$, we have $\tau_1(k,l) = T + k_2$. Therefore, $N_1 = n$ if and only if $n = 0$. In other words, $\mathbb{P}_x(N_1 = n \mid Y_{T+1} = 2^{k_1}-1, \ldots, Y_{T+k_2-1} = 2^{k_1}-1, Y_{T+k_2} = 2^{k_1}-1) = \mathbb{I}_n(0)$ where $\mathbb{I}$ is the indicator function as defined in the statement of the Theorem.

Thus combining the above expression with the equations (16) - (22), we have

$$g_1^{(x)}(n) = q_{2^m-1}g_1^{(2^m-2)}(n-1) + \sum_{t=1}^{k_2-1} q_{2^m-1}\left(p_{2^m-1}\right)^t g_1^{(2^m-2)}(n) + \left(p_{2^m-1}\right)^{k_2}\mathbb{I}_n(0).$$

This completes the proof. □

We note that the right hand side of (15) does not involve the initial condition $x \in C_m$. Therefore $g_1^{(x)}(n)$ must be independent of x. So, we will drop x and denote the above probability by $g_1(n)$, *i.e.*,

$$g_1(n) = \mathbb{P}_x(N_1 = n).$$

Hence, we may rewrite the equation (15) as follows: for any $k_2 = k - k_1 > 0$, $x \in C_{k_1}$ and any $n \geq 0$,

$$g_1(n) = q_{2^m-1} g_1(n-1) + \sum_{t=1}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^t g_1(n) + \Big(p_{2^m-1}\Big)^{k_2} \mathbb{I}_n(0). \tag{23}$$

Now, the equation (23) can be easily solved.

Corollary 2: Suppose that $k_2 = k - k_1 > 0$. For any $x \in C_{k_1}$ and any $n \geq 0$, we have

$$g_1(n) = \left[\frac{q_{2^m-1}}{1 - \sum_{t=1}^{k_2-1}\Big(p_{2^m-1}\Big)^t q_{2^m-1}}\right]^n \frac{\Big(p_{2^m-1}\Big)^{k_2}}{1 - \sum_{t=1}^{k_2-1}\Big(p_{2^m-1}\Big)^t q_{2^m-1}}. \tag{24}$$

Indeed, for $n = 0$, we have

$$g_1(0) = \frac{\Big(p_{2^m-1}\Big)^{k_2}}{1 - \sum_{t=1}^{k_2-1}\Big(p_{2^m-1}\Big)^t q_{2^m-1}} = \frac{\Big(p_{2^m-1}\Big)^{k_2}}{q_{2^m-1} + \Big(p_{2^m-1}\Big)^{k_2}}. \tag{25}$$

For $n \geq 1$, inductively we have

$$g_1(n) = \frac{g_1(n-1)\ q_{2^m-1}}{1 - \sum_{t=1}^{k_2-1}\Big(p_{2^m-1}\Big)^t q_{2^m-1}} = \left[\frac{q_{2^m-1}}{1 - \sum_{t=1}^{k_2-1}\Big(p_{2^m-1}\Big)^t q_{2^m-1}}\right]^n g_1(0)$$

which proves the corollary.

We observe that N_1 follows a geometric distribution with parameter p_E where p_E is given in (7). The generating function of N_1 is given by

$$\zeta_1(s; k_1) = \frac{p_E}{1 - (1 - p_E)s} = \frac{\Big(p_{2^m-1}\Big)^{k_2}}{q_{2^m-1} + \Big(p_{2^m-1}\Big)^{k_2} - q_{2^m-1}s}. \tag{26}$$

For $r \geq 2$, we can also derive a similar recurrence relation.

Theorem 4: Suppose that $k_2 = k - k_1 > 0$. For any $x \in C_{k_1}$ and any $n \geq 0, r \geq 2$, we have

$$\begin{aligned} g_r^{(x)}(n) &= q_{2^m-1} g_r^{(2^m-2)}(n-1) + \sum_{t=1}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^t g_r^{(2^m-2)}(n) \\ &+ \sum_{j_1=0}^{r-2}\sum_{j_2=0}^{l} q_{2^m-1}\Big(p_{2^m-1}\Big)^{k_2+j_1(l+1)+j_2} g_{r-1-j_1}^{(2^m-2)}(n) + \Big(p_{2^m-1}\Big)^{k_2+(r-1)(l+1)} \mathbb{I}_n(0). \end{aligned} \tag{27}$$

where $\mathbb{I}$ is the indicator function as defined earlier.

The proof is very similar to the proof of Theorem 3. Again, conditioning on the process when T occurs, we obtain for $k_2 > 0$, as in Theorem 3,

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

The above expression is similar to the expression given in (16) obtained in Theorem 3. Hence following the similar calculations, we get the required result in Theorem 4.

The recurrence relation in (27) cannot be solved directly. However, we may easily check that $g_r^{(x)}(\cdot)$ is independent of x. We have shown that $g_r^{(x)}(\cdot)$ is independent of x for $r = 1$. By induction, assume that $g_r^{(x)}(\cdot)$ is independent of $x \in C_m$. Clearly, from the relation (27), we have that $g_{r+1}^{(x)}(\cdot)$ can be expressed as weighted sums of $g_i^{(x)}(\cdot)$ for $i = 1, 2, \ldots, r$ and other terms which do not involve x. Since the right hand side of the above relation does not involve any $x \in C_m$, the left hand side, *i.e.*, $g_{r+1}^{(x)}(\cdot)$ must be independent of x. Therefore, from now on, we will drop the superscript x from the notation and denote it by $g_r(\cdot)$.

The equation in (27) may now be simplified. Transferring terms containing $g_r(n)$ in the right hand side to the left hand side, we have the following result.

Lemma 1: Suppose that $k_2 = k - k_1 > 0$. For any $x \in C_{k_1}$ and any $n \geq 0, r \geq 1$, $g_r^{(x)}(n) = \mathbb{P}_x(N_r = n)$ is independent of x. For $r \geq 2$, it satisfies the recurrence relation

$$\begin{aligned} &\Big(1 - \sum_{j=1}^{k_2-1} q_{2^m-1}\big(p_{2^m-1}\big)^j\Big) g_r(n) \\ &= q_{2^m-1} g_r(n-1) + \big(p_{2^m-1}\big)^{k_2}\Big(1 - \big(p_{2^m-1}\big)^{l+1}\Big)\sum_{j=0}^{r-2}\big(p_{2^m-1}\big)^{j(l+1)} g_{r-1-j}(n) \\ &\qquad + \big(p_{2^m-1}\big)^{k_2+(r-1)(l+1)} \mathbb{I}_n(0). \end{aligned} \tag{28}$$

Now, using relation (28), we develop the recurrence relation between the probability generating functions of N_r. The probability generating function $\zeta_r(s; k_1)$, for $r \geq 2$ and $k_2 > 0$, is given by

$$\begin{aligned} &\Big(1 - \sum_{j=1}^{k_2-1} q_{2^m-1}\big(p_{2^m-1}\big)^j\Big)\zeta_r(s; k_1) = \sum_{n=0}^{\infty}\Big(1 - \sum_{j=1}^{k_2-1} q_{2^m-1}\big(p_{2^m-1}\big)^j\Big) g_r(n) s^n \\ &= \big(p_{2^m-1}\big)^{k_2+(r-1)(l+1)} + \sum_{n=0}^{\infty} q_{2^m-1} g_r(n-1) s^n \end{aligned}$$

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

Thus, we have proved the following lemma.

Lemma 2: For $r \geq 2$ and $k_2 > 0$, the sequence of probability generating functions satisfy the recurrence relation

$$
\begin{aligned}
&\Big(1 - q_{2^m-1}s - \sum_{j=1}^{k_2-1} q_{2^m-1}\Big(p_{2^m-1}\Big)^{j}\Big)\zeta_r(s;k_1) \\
&= \Big(p_{2^m-1}\Big)^{k_2}\Big(1-\Big(p_{2^m-1}\Big)^{l+1}\Big)\sum_{j=0}^{r-2}\Big(p_{2^m-1}\Big)^{j(l+1)}\zeta_{r-1-j}(s;k_1) \\
&+ \Big(p_{2^m-1}\Big)^{k_2+(r-1)(l+1)}. \qquad (29)
\end{aligned}
$$

Now we can use the above results to prove the main Theorem 1 as follows:

Proof: (**Theorem 1**) Let the generating function of the sequence $\{\zeta_r(s;k_1) : r \geq 1\}$ be denoted by $\Xi(z;k_1)$, *i.e.*, $\Xi(z;k_1) = \sum_{r=1}^{\infty} \zeta_r(s;k_1)z^r$. For $k_2 > 0$, we have

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

$$= \frac{\left(p_{2^m-1}\right)^{k_2} z}{1-\left(p_{2^m-1}\right)^{(l+1)} z} + z\Xi(z;k_1)\left(p_{2^m-1}\right)^{k_2}\left(1-\left(p_{2^m-1}\right)^{l+1}\right)\sum_{j=0}^{\infty}\left(p_{2^m-1}\right)^{j(l+1)} z^j$$

$$= \frac{\left(p_{2^m-1}\right)^{k_2} z}{1-\left(p_{2^m-1}\right)^{(l+1)} z} + \frac{z\Xi(z;k_1)\left(p_{2^m-1}\right)^{k_2}\left(1-\left(p_{2^m-1}\right)^{l+1}\right)}{1-\left(p_{2^m-1}\right)^{(l+1)} z}. \tag{30}$$

Now, from the above equation (30), we can easily solve for $\Xi(z;k_1)$ to obtain

$$\begin{aligned}
&\Xi(z;k_1) \\
&= \left[\left(p_{2^m-1}\right)^{k_2} z\right]\Bigg[\left(1-\left(p_{2^m-1}\right)^{l+1} z\right)\left(1-q_{2^m-1}s-\sum_{j=1}^{k_2-1} q_{2^m-1}\left(p_{2^m-1}\right)^j\right) \\
&\qquad - z\left(p_{2^m-1}\right)^{k_2}\left(1-\left(p_{2^m-1}\right)^{l+1}\right)\Bigg]^{-1} \\
&= \frac{z\left(p_{2^m-1}\right)^{k_2}}{1-q_{2^m-1}s-\sum_{j=1}^{k_2-1} q_{2^m-1}\left(p_{2^m-1}\right)^j} \\
&\qquad \times\left[1-\left(p_{2^m-1}\right)^{l+1} z - \frac{z\left(p_{2^m-1}\right)^{k_2}\left(1-\left(p_{2^m-1}\right)^{l+1}\right)}{1-q_{2^m-1}s-\sum_{j=1}^{k_2-1} q_{2^m-1}\left(p_{2^m-1}\right)^j}\right]^{-1} \\
&= \frac{z\left(p_{2^m-1}\right)^{k_2}}{1-q_{2^m-1}s-\sum_{j=1}^{k_2-1} q_{2^m-1}\left(p_{2^m-1}\right)^j} \\
&\qquad \times\left[1-z\left[\left(p_{2^m-1}\right)^{l+1} + \frac{\left(p_{2^m-1}\right)^{k_2}\left(1-\left(p_{2^m-1}\right)^{l+1}\right)}{1-q_{2^m-1}s-\sum_{j=1}^{k_2-1} q_{2^m-1}\left(p_{2^m-1}\right)^j}\right]\right]^{-1}.
\end{aligned} \tag{31}$$

From the expression of generating function $\Xi(z;k_1)$, $\zeta_r(s;k_1)$ is obtained by computing the coefficient of z^r. Observe that first term in the right hand side of (31) has a power of z. Therefore, we need a power of z^{r-1} from the second term. Using the expansion $(1-az)^{-1} = \sum_{n=0}^{\infty} a^n z^n$, we have

$$\begin{aligned}
\zeta_r(s;k_1) = &\left[\frac{\left(p_{2^m-1}\right)^{k_2}}{q_{2^m-1}+\left(p_{2^m-1}\right)^{k_2}-q_{2^m-1}s}\right]\Bigg[\left(p_{2^m-1}\right)^{l+1} \\
&+ \frac{\left(p_{2^m-1}\right)^{k_2}}{q_{2^m-1}+\left(p_{2^m-1}\right)^{k_2}-q_{2^m-1}s}\left(1-\left(p_{2^m-1}\right)^{l+1}\right)\Bigg]^{r-1}.
\end{aligned} \tag{32}$$

This completes the proof. □

If we consider the case, when $k_2 = 0$, *i.e.*, $k = k_1$, then for $r = 1$, we must have

$\mathbb{P}_x(N_1(k) = 1) = 1$. Thus the probability generating function is given by

$$\zeta_1(s;k) = s$$

However for $r \geq 2$, using the similar arguments, we obtain,

$$g_r(n) = q_{2^m-1} g_{r-1}(n-1) + \Big(1 - q_{2^m-1} - \big(p_{2^m-1}\big)^{l+1}\Big) g_{r-1}(n) + \sum_{j=1}^{r-2} \big(p_{2^m-1}\big)^{j(l+1)} \Big(1 - \big(p_{2^m-1}\big)^{l+1}\Big) g_{r-1-j}(n) + \big(p_{2^m-1}\big)^{(r-1)(l+1)} \mathbb{I}_n(0).$$

This again can be used to obtain the recurrence relation between the probability generating functions $\zeta_r(s;k)$ for $r \geq 2$. Indeed, we would obtain

$$\zeta_r(s;k) = \big(p_{2^m-1}\big)^{(r-1)(l+1)} + q_{2^m-1} s \zeta_{r-1}(s;k) + \Big(1 - q_{2^m-1} - \big(p_{2^m-1}\big)^{l+1}\Big) \zeta_{r-1}(s;k) + \Big(1 - \big(p_{2^m-1}\big)^{l+1}\Big) \sum_{j=1}^{r-2} \big(p_{2^m-1}\big)^{j(l+1)} \zeta_{r-1-j}(s;k).$$

Using the above expression, we obtain the generating function $\Xi(z;k)$ as follows:

$$\Xi(z;k) = \frac{sz + \big(p_{2^m-1}\big)^{l+1}(1-s)}{1 - z(p_{2^m-1} + q_{2^m-1}s) - z^2 \big(p_{2^m-1}\big)^{l+1}(1 - p_{2^m-1} - q_{2^m-1}s)}. \tag{33}$$

However, the explicit expression for $\zeta_r(s;k)$, *i.e.*, the coefficient of z^r in (33), will be complicated and it would be difficult to identify the distribution of the underlying random variables in terms of the known probability distributions..

5. Strong law of large numbers

In this section, we show how we may use our main result to establish the strong law of large numbers for the number of runs of exact length. Given k_1, we may fix $k = k_1 + 1$. For simplicity of the calculations, we will consider here the non-overlapping runs, *i.e.*, $l = k - 1 = k_1$. Let us define

$$\theta(n) = \sup\{r \geq 0 : \tau_r(k_1 + 1, k_1) \leq n\}. \tag{34}$$

Clearly, $\theta(n)$ represents the number of non-overlapping runs of length k that have been observed until time n. Also, we must have

$$\tau_{\theta(n)}(k_1 + 1, k_1) \leq n < \tau_{\theta(n)+1}(k_1 + 1, k_1).$$

First we observe that the occurrence of a non-overlapping run is a renewal event in our set up. Let E_t denote the event that a non-overlapping run has finished at time t. Then, for $t, s \geq 1$, we have

$$\mathbb{P}_x(E_t \cap E_{t+s}) = \mathbb{P}_x(E_t)\mathbb{P}_x\Big(E_{t+s} \mid (Y_u; u \leq t)\Big)$$

$$= \mathbb{P}_x(E_t)\mathbb{P}_x\Big(E_{t+s} \mid Y_t = 2^{k_1-1}\Big)$$
$$= \mathbb{P}_x(E_t)\mathbb{P}_{(2^m-1)}\Big(E_s\Big)$$

where we have used the strong Markov property on the expression in second step and the fact that at time t, a non-overlapping run is finished and hence we must have $Y_t = 2^{k_1-1}$. Further, this shows that the events again have the structure of a delayed renewal event.

Since we have assumed that $0 < p_x < 1$, for all $x \in C_i$, it is the case that the Markov chain $\{Y_t : t \geq 0\}$ is an irreducible chain and hence positive recurrent. This implies that the renewal event is also positive recurrent. Therefore, the expected time for getting $k_1 + 1$ consecutive successes from any state is finite and have finite expectation. In other words, we must have

$$\mathbb{E}_{(2^m-1)}(\tau_1(k_1+1, k_1)) = \mu < \infty. \tag{35}$$

The value of μ will depend upon the values of $\{p_x : x \in C_i\}$. For the *i.i.d.* case, it is known that (see Feller (1968), page 324),

$$\mu = \frac{1 - p^{k_1+1}}{qp^{k_1+1}}.$$

Using the results of renewal theory (see Feller (1968)), we further have that

$$\frac{1}{n}\theta(n) \to \frac{1}{\mu} \tag{36}$$

with probability 1. Now, we prove Theorem 2 which establishes the strong law of large numbers.

Proof: (**Theorem 2**) For any $r \geq 1$, we can represent, using Corollary 1,

$$N_{\tau_r(k_1+1,k_1)}(k_1)(= N_r(k_1)) \stackrel{d}{=} G_1 + \sum_{i=2}^{r} G_i B_i.$$

Since the equality is in distribution, we cannot directly apply the strong law on this family to conclude our result.

Now, expectation of the random variable G_1 as well as G_1B_1 may be computed from the probability generating function given in equation (9). Indeed, we have $\mathbb{E}(G_1B_1) = \mu_1$ (see equation (10)). Further observe that all moments of G_1B_1 are finite.

Let us set $\mu_1(r) = \Big[\mathbb{E}(G_1) + (r-1)\mathbb{E}(G_1B_1)\Big]$. Then, we have

$$\frac{1}{r}\mu_1(r) = \frac{1}{r}\Big[\mathbb{E}(G_1) + (r-1)\mu_1\Big] \to \mu_1 \tag{37}$$

as $r \to \infty$. Note that, from the representation, we have $\mathbb{E}(N_{\tau_r(k_1+1,k_1)}(k_1)) = \mu_1(r)$. Furthermore, for any $\epsilon > 0$, we have

$$\mathbb{P}\Big(\frac{1}{r}|N_{\tau_r(k_1+1,k_1)}(k_1) - \mu_1(r)| \geq \epsilon\Big) = \mathbb{P}\Big(\frac{1}{r}|G_1 + \sum_{i=2}^{r} G_i B_i - \mu_1(r)| \geq \epsilon\Big)$$

$$= \mathbb{P}\Big(|G_1 - \mathbb{E}(G_1) + \sum_{i=2}^{r}(G_iB_i - \mathbb{E}(G_1B_1))| \geq r\epsilon\Big).$$

Now, we may estimate the probability using the Markov inequality. Indeed, we have

$$\begin{aligned}
&\mathbb{P}\Big[|G_1 - \mathbb{E}(G_1) + \sum_{i=2}^{r} G_iB_i - \mathbb{E}(G_1B_1)| \geq r\epsilon\Big] \\
&\leq \frac{1}{r^4\epsilon^4}\mathbb{E}\Big[\Big(G_1 - \mathbb{E}(G_1) + \sum_{i=2}^{r} G_iB_i - \mathbb{E}(G_1B_1)\Big)^4\Big] \\
&\leq \frac{1}{r^4\epsilon^4}\Big[\mathbb{E}\Big(G_1 - \mathbb{E}(G_1)\Big)^4 + 3(r-1)\mathbb{E}\Big(G_1 - \mathbb{E}(G_1)\Big)^2 E\Big(G_1B_1 - \mathbb{E}(G_1B_1)\Big)^2 \\
&\qquad + 6(r-1)^2\Big(E\Big(G_1B_1 - \mathbb{E}(G_1B_1)\Big)^2\Big)^2 + (r-1)\mathbb{E}\Big(G_1B_1 - \mathbb{E}(G_1B_1)\Big)^4\Big] \\
&\leq \frac{C}{r^2\epsilon^4}
\end{aligned}$$

for a suitably chosen constant $C > 0$.

Thus, by Borel-Cantelli lemma, we conclude that $\frac{1}{r}\Big(N_{\tau_r(k_1+1,k_1)}(k_1) - \mu_1(r)\Big) \to 0$ with probability 1. This along with equation (37) implies that

$$\frac{1}{r}N_{\tau_r(k_1+1,k_1)}(k_1) \to \mu_1 \tag{38}$$

as $r \to \infty$ with probability 1.

Since $\tau_{\theta(n)}(k_1+1, k_1) \leq n < \tau_{\theta(n)+1}(k_1+1, k_1)$, we must have $N_{\theta(n)}(k_1) \leq N_n(k_1) \leq N_{\theta(n)+1}(k_1)$. Therefore, we obtain that

$$\begin{aligned}
&\frac{1}{n}N_{\theta(n)}(k_1) \leq \frac{1}{n}N_n(k_1) \leq \frac{1}{n}N_{\theta(n)+1}(k_1) \\
&\implies \frac{\theta(n)}{n} \times \frac{1}{\theta(n)}N_{\theta(n)}(k_1) \leq \frac{1}{n}N_n(k_1) \leq \frac{\theta(n)+1}{n} \times \frac{1}{\theta(n)+1}N_{\theta(n)+1}(k_1).
\end{aligned}$$

Since $\theta(n)/n \to \frac{1}{\mu}$, we have that $\theta(n) \to \infty$ as $n \to \infty$. Hence, we may apply the equation (38) along the sub-sequence $\theta(n)$ and equation (36) to conclude that both the upper bound as well as the lower bound will converge to μ_1/μ. This proves the result. □

Acknowledgement

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

References

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). Numbers 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. (2022). Asymptotic results for generalized runs in higher order Markov Chains. *Statistics and Applications*, **21**, 189-207.

Balakrishnan, N. and Koutras, M. V. (2002). *Runs and Scans with Applications.* John Wiley and Sons, New York.

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.

Doganaksoy, A., Sulak, F., Uguz, M., Seker, O., and Akcengiz, Z. (2015). New statistical randomness tests based on length of runs. *Mathematical Problems in Engineering*, **2015**, 1-14.

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

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. *Advances in Combinatorial Methods and Applications to Probability and Statistics,* 401–410, Statistics for Industry and Technology, Birkhäuser Boston, Boston, MA.

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.