Ann. Inst. Statist. Math.
Vol. 50, No. 4, 655–671 (1998)

ON GENERATING FUNCTIONS OF WAITING TIME PROBLEMS FOR SEQUENCE PATTERNS OF DISCRETE RANDOM VARIABLES*

MASAYUKI UCHIDA

The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan

(Received July 31, 1996; revised August 28, 1997)

Abstract. Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, which take values in a countable set $S = \{0, 1, 2, \ldots\}$. By a pattern we mean a finite sequence of elements in S. For every $i = 0, 1, 2, \ldots,$ we denote by $P_i =$ "$a_{i,1}a_{i,2}\cdots a_{i,k_i}$" the pattern of some length k_i, and E_i denotes the event that the pattern P_i occurs in the sequence $X_1, X_2, \ldots$. In this paper, we have derived the generalized probability generating functions of the distributions of the waiting times until the r-th occurrence among the events $\{E_i\}_{i=0}^{\infty}$. We also have derived the probability generating functions of the distributions of the number of occurrences of sub-patterns of length $l (l < k)$ until the first occurrence of the pattern of length k in the higher order Markov chain.

Key words and phrases: Sequence patterns, runs, sooner and later problems, r-th occurrence problem, discrete distribution of order k, generalized probability generating function, higher order Markov chain.

1. Introduction

In recent years exact discrete distribution theory related to succession events has been developed. Typical distributions are called discrete distributions of order k. For example, Philippou *et al.* (1983) introduced the geometric distribution of order k. The geometric distribution of order k is the distribution of the number of trials until the first occurrence of the k-th consecutive succession in a sequence of independent Bernoulli trials with common success probability p. There are some applications of the geometric distribution of order k, for example, the evaluation of a start-up demonstration test (Hahn and Gage (1983) and Viveros and Balakrishnan (1993)). The geometric distribution of order k is one of the simplest waiting time distributions. Several waiting time problems have been studied

* The research was supported in part by Grant-in-Aid for Encouragement of Young Scientists from the Ministry of Education, Science and Culture.

by many authors in more general situations (Li (1980), Gerber and Li (1981), Ebneshahrashoob and Sobel (1990), Ling (1990), Aki (1992), Aki and Hirano (1993), Balasubramanian *et al.* (1993), Fu (1996), Fu and Koutras (1994), Uchida and Aki (1995) and references therein).

An interesting class of waiting time problems was proposed by Ebneshahrashoob and Sobel (1990). They obtained the generalized probability generating functions of the waiting time distributions for a run of "0" of length r or (and) a run of "1" of length k whichever comes sooner (later) when the sequence $X_1, X_2, \ldots$ is constructed from Bernoulli trials, that is, X's are i.i.d. and $\{0, 1\}$-valued random variables. Uchida and Aki (1995) applied the sooner waiting time problem to the volleyball game.

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, which take values in a countable set $S = \{0, 1, 2, \ldots\}$. By a *pattern* it means a finite sequence of elements in S. Suppose that a sequence of positive integers $\{k_i\}_{i=0}^{\infty}$ is given. Let P_i = "$a_{i,1}a_{i,2}\cdots a_{i,k_i}$" be a pattern. Then P_i is the pattern of length k_i, where k_i is number of its elements in the string. Let $\{P_i\}$ be a set generated by all the elements of the pattern P_i, that is, $\{P_i\} = \{a_{i,1}, a_{i,2}, \ldots, a_{i,k_i}\}$. In this paper, we assume that $\{P_i\} \bigcap \{P_j\} = \phi$ for $i \neq j$ and $a_{l,1} \leq a_{l,2} \leq \cdots \leq a_{l,k_l}$ for $l = 0, 1, 2, \ldots$. This assumption is not so strong because if P_0 does not satisfy the above condition we can partition P_0 into several subpatterns so that increasing order can be maintained in each subpattern. For every $i = 0, 1, 2, \ldots$, we denote by E_i the event that the pattern P_i occurs in the sequence $X_1, X_2, \ldots$.

In Section 2, we firstly obtain the probability generating function of the distribution of the waiting time for the occurrence of the event E_0 in the sequence $X_1, X_2, \ldots$. Next, we have the generalized probability generating function of the distribution of the waiting time for the sooner occurring event between E_0 and E_1 in the sequence $X_1, X_2, \ldots$. Finally, we derive the probability generating function of the distribution of the waiting time for the later occurring event between E_0 and E_1 in the sequence $X_1, X_2, \ldots$. In Section 3, we obtain the generalized probability generating function of the distribution of the waiting time until at least one of the events $\{E_i\}_{i=0}^{\infty}$ occurs. Next, we have the generalized probability generating function of the distribution of the waiting time for the second occurrence among $\{E_i\}_{i=0}^{\infty}$. Here, "the second occurrence" means the occurrence of another event excepting the first event among the events $\{E_i\}_{i=0}^{\infty}$. In general, the generalized probability generating function of the distribution of the waiting time for the r-th occurrence of the event among $\{E_i\}_{i=0}^{\infty}$ is considered. Aki (1992) solved this type of problem when E_i, for every $i = 0, 1, 2, \ldots$, is the event that a run of "i" of length k_i occurs in the sequence $X_1, X_2, \ldots$, that is, the waiting time problem for the r-th occurrence of the event among $\{E_i\}_{i=0}^{\infty}$. However, in this paper, the probability distribution of waiting time of r-th occurrence of a compound pattern could not be obtained from our generating functions.

For the derivation of the main part of the results, we use the method of generalized probability generating function (Ebneshahrashoob and Sobel (1990)).

Aki and Hirano (1994) showed that, when $\{0, 1\}$-sequence follows the first order Markov chain, the distribution of overlapping occurrences of success-runs of length l until the first consecutive k successes is the shifted geometric distribution of order $k-l$ with the support $\{k-l+1, k-l+2, \ldots\}$, where we usually regard the valued 1 as success and the value 0 as failure. Further, Aki and Hirano (1995) extended the result and studied the joint distributions of the numbers of trials and of outcomes such as successes, failures and success-runs until the first consecutive k successes in the first order Markov dependent trials.

Now suppose that $\{0, 1, 2, \ldots\}$-sequence follows the m-th order Markov chain. Let P_0 = "$a_1 a_2 \cdots a_k$" be a pattern of length k, where $a_1 \leq a_2 \leq \cdots \leq a_k$. A pattern S_l = "$a_1 a_2 \cdots a_l$" is called a *sub-pattern* of the pattern P_0 for $l = 1, 2, \ldots, k-1$. In Section 4, we study the distributions of number of occurrences of sub-patterns of length l, that is, S_l until the first occurrence of the pattern P_0 of length k. Throughout the last section, let k, l and m be fixed positive integers such that $m \leq l < k$. We denote by $G_k(p)$ the geometric distribution of order k, and by $G_k(p, a)$ the shifted geometric distribution of order k so that its support begins with a.

2. Waiting time problems for two patterns

2.1 *Waiting time problem for a single pattern*

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, which take values in a countable set $S = \{0, 1, 2, \ldots\}$. Let $p_i = P\{X_1 = i\}$, $i = 0, 1, 2, \ldots$. Let P_0 = "$a_1 a_2 \cdots a_k$" be a pattern of length k, where $a_1 \leq a_2 \leq \cdots \leq a_k$. Let E_0 be the event that the pattern P_0 occurs in the sequence $X_1, X_2, \ldots$. In this subsection, we consider the distribution of the waiting time for the occurrence of the event E_0. Let τ be the waiting time for the occurrence of the event E_0. Let $\varphi(t)$ be the p.g.f. of the distribution of the waiting time τ. Let t_0 be an arbitrary positive integer and let $\varphi^i(t)$ be the p.g.f. of the conditional distribution of $\tau - t_0$ given that the pattern "$a_1 a_2 \cdots a_i$" has just occurred at the t_0-th trial, where $i = 1, 2, \ldots, k$ and $\varphi^k(t) = 1$. We also note that $\varphi^i(t)$, $i = 1, 2, \ldots, k$ does not depend on t_0. Then φ and φ^i, $i = 1, 2, \ldots, k$ satisfy the following system of linear equations:

$$(2.1)\quad \varphi = p_{a_1} t\varphi^1 + (1 - p_{a_1})t\varphi,$$

$$(2.2)\quad \varphi^i = p_{a_{i+1}} t\varphi^{i+1} + p_{a_1} 1\{a_{i+1} \neq a_1\}1\{a_i \neq a_1\}t\varphi^1 + p_{a_1} 1\{a_{i+1} \neq a_1\}1\{a_i = a_1\}t\varphi^i + (1 - p_{a_{i+1}} - p_{a_1} 1\{a_{i+1} \neq a_1\})t\varphi \quad \text{for } i = 1, 2, \ldots, k-1,$$

where for $i = 1, 2, \ldots, k$,

$$1\{a_i \neq a_1\} = \begin{cases} 1, & \text{if} \quad a_i \neq a_1, \\ 0, & \text{if} \quad a_i = a_1. \end{cases}$$

We set that for $i = 1, 2, \ldots, k-1$,

$$A_i = p_{a_1} 1\{a_{i+1} \neq a_1\}1\{a_i \neq a_1\},$$
$$B_i = p_{a_1} 1\{a_{i+1} \neq a_1\}1\{a_i = a_1\}.$$

From (2.2), we obtain for $i = 1, 2, \dots, k-1$,

$$(1 - B_i t)\varphi^i = p_{a_{i+1}} t\varphi^{i+1} + A_i t\varphi^1 + (1 - p_{a_{i+1}} - A_i - B_i)t\varphi. \tag{2.3}$$

Here, we set that for $i = 1, 2, \dots, k-1$,

$$\alpha_i = \frac{p_{a_{i+1}} t}{1 - B_i t},$$
$$\beta_i = \frac{A_i t}{1 - B_i t},$$
$$\gamma_i = \frac{(1 - p_{a_{i+1}} - A_i - B_i)t}{1 - B_i t}.$$

From (2.3), we obtain for $i = 1, 2, \dots, k-1$,

$$\varphi^i = \alpha_i \varphi^{i+1} + \beta_i \varphi^1 + \gamma_i \varphi.$$

So, we have

$$\varphi^1 = \left[\prod_{i=1}^{k-1} \alpha_i\right] + \sum_{i=1}^{k-1} \left[\prod_{j=1}^{i-1} \alpha_j\right] \beta_i \varphi^1 + \sum_{i=1}^{k-1} \left[\prod_{j=1}^{i-1} \alpha_j\right] \gamma_i \varphi, \tag{2.4}$$

where we define that $\prod_{j=1}^{0} \alpha_j = 1$.

From (2.1) and (2.4), we obtain

$$\begin{aligned}
[1 - (1 - p_{a_1})t]\varphi &= p_{a_1} t \left[\prod_{i=1}^{k-1} \alpha_i\right] + \sum_{i=1}^{k-1} \left[\prod_{j=1}^{i-1} \alpha_j\right] \beta_i p_{a_1} t\varphi^1 + \sum_{i=1}^{k-1} \left[\prod_{j=1}^{i-1} \alpha_j\right] p_{a_1} t\gamma_i \varphi \\
&= p_{a_1} t \left[\prod_{i=1}^{k-1} \alpha_i\right] + \sum_{i=1}^{k-1} \left[\prod_{j=1}^{i-1} \alpha_j\right] \beta_i [1 - (1 - p_{a_1})t]\varphi \\
&\quad + \sum_{i=1}^{k-1} \left[\prod_{j=1}^{i-1} \alpha_j\right] p_{a_1} t\gamma_i \varphi \\
&= p_{a_1} t \left[\prod_{i=1}^{k-1} \alpha_i\right] + \sum_{i=1}^{k-1} \left[\prod_{j=1}^{i-1} \alpha_j\right] (\beta_i [1 - (1 - p_{a_1})t] + p_{a_1} t\gamma_i)\varphi \\
&= p_{a_1} t \left[\prod_{i=1}^{k-1} \alpha_i\right] \\
&\quad + \sum_{i=1}^{k-1} \left[\prod_{j=1}^{i-1} \alpha_j\right] \left[-(1 - t)\frac{p_{a_1} t - A_i t}{1 - B_i t} + p_{a_1} t - p_{a_1} t\alpha_i\right] \varphi \\
&= p_{a_1} t \left[\prod_{i=1}^{k-1} \alpha_i\right] \\
&\quad + \left[-(1 - t)\sum_{i=1}^{k-1} \frac{p_{a_1} t - A_i t}{1 - B_i t} \left(\prod_{j=1}^{i-1} \alpha_j\right) + p_{a_1} t - p_{a_1} t \left(\prod_{i=1}^{k-1} \alpha_i\right)\right] \varphi.
\end{aligned}$$

Hence, we have

LEMMA 2.1. *The p.g.f.* $\varphi(t)$ *is given by*

$$\varphi(t) = \frac{\Gamma(t)}{1 - t + \Gamma(t)},$$

where

$$\Gamma(t) = \frac{p_{a_1} t (\prod_{i=1}^{k-1} \alpha_i)}{1 + \sum_{i=1}^{k-1} \frac{p_{a_1} t - A_i t}{1 - B_i t} (\prod_{j=1}^{i-1} \alpha_j)}.$$

2.2 *Sooner waiting time problem for two patterns*

Let $P_0 =$ "$a_1 a_2 \cdots a_k$" be a pattern of length k and let $P_1 =$ "$b_1 b_2 \cdots b_l$" be a pattern of length l. We assume that $\{P_0\} \bigcap \{P_1\} = \phi$ and $a_1 \leq a_2 \leq \cdots \leq a_k$, $b_1 \leq b_2 \leq \cdots \leq b_l$. Let E_0 be the event that the pattern P_0 occurs and let E_1 be the event that the pattern P_1 occurs in the sequence $X_1, X_2, \ldots$. In this subsection, we consider the distribution of the waiting time for the sooner occurring event between E_0 and E_1. We derive a generalized probability generating function (g.p.g.f.) by adding markers x_i, $i = 0, 1$. Here, x_i represents that the sooner occurring event is E_i. Let τ be the waiting time for the sooner occurring event between E_0 and E_1. Let $\phi(t)$ be the g.p.g.f. of the distribution of the waiting time τ. Let t_0 be an arbitrary positive integer and let $\varphi^i(t)$ be the g.p.g.f. of the conditional distribution of $\tau - t_0$ given that the pattern "$a_1 a_2 \cdots a_i$" has just occurred at the t_0-th trial, where $i = 1, 2, \ldots, k$ and $\varphi^k(t) = x_0$. Let $\varphi_j(t)$ be the g.p.g.f. of the conditional distribution of $\tau - t_0$ given that the pattern "$b_1 b_2 \cdots b_j$" has just occurred at the t_0-th trial, where $j = 1, 2, \ldots, l$ and $\varphi_l(t) = x_1$. We also note that $\varphi^i(t)$ and $\varphi_j(t)$, $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, l$ do not depend on t_0.

Here, we set that for $i = 1, 2, \ldots, k - 1$,

$$\begin{aligned} A_i &= p_{a_1} 1\{a_{i+1} \neq a_1\} 1\{a_i \neq a_1\}, \\ B_i &= p_{a_1} 1\{a_{i+1} \neq a_1\} 1\{a_i = a_1\}, \end{aligned}$$

and for $j = 1, 2, \ldots, l - 1$,

$$\begin{aligned} C_j &= p_{b_1} 1\{b_{j+1} \neq b_1\} 1\{b_j \neq b_1\}, \\ D_j &= p_{b_1} 1\{b_{j+1} \neq b_1\} 1\{b_j = b_1\}. \end{aligned}$$

Then φ, φ^i and φ_j, $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, l$ satisfy the following system of linear equations:

$$\varphi = p_{a_1} t \varphi^1 + p_{b_1} t \varphi_1 + (1 - p_{a_1} - p_{b_1}) t \varphi, \tag{2.5}$$

$$\begin{aligned} \varphi^i &= p_{a_{i+1}} t \varphi^{i+1} + A_i t \varphi^1 \\ &\quad + B_i t \varphi^i + p_{b_1} t \varphi_1 + (1 - p_{a_{i+1}} - A_i - B_i - p_{b_1}) t \varphi, \\ &\qquad \text{for} \quad i = 1, 2, \ldots, k - 1, \end{aligned} \tag{2.6}$$

$$\begin{aligned} \varphi_j &= p_{b_{j+1}} t \varphi_{j+1} + C_j t \varphi_1 \\ &\quad + D_j t \varphi_j + p_{a_1} t \varphi^1 + (1 - p_{b_{j+1}} - C_j - D_j - p_{a_1}) t \varphi, \end{aligned} \tag{2.7}$$

$$\text{for} \quad j = 1, 2, \ldots, l-1.$$

We set that for $i = 1, 2, \ldots, k-1$,

$$\alpha_i = \frac{p_{a_{i+1}} t}{1 - B_i t},$$
$$\beta_i = \frac{A_i t}{1 - B_i t},$$
$$\gamma_i = \frac{1}{1 - B_i t},$$
$$\delta_i = \frac{(1 - p_{a_{i+1}} - A_i - B_i - p_{b_1}) t}{1 - B_i t},$$

and for $j = 1, 2, \ldots, l-1$,

$$\eta_j = \frac{p_{b_{j+1}} t}{1 - D_j t},$$
$$\mu_j = \frac{C_j t}{1 - D_j t},$$
$$\nu_j = \frac{1}{1 - D_j t},$$
$$\xi_j = \frac{(1 - p_{b_{j+1}} - C_j - D_j - p_{a_1}) t}{1 - D_j t}.$$

From (2.6) and (2.7), we have

$$\varphi^i = \alpha_i \varphi^{i+1} + \beta_i \varphi^1 + p_{b_1} t \gamma_i \varphi_1 + \delta_i \varphi, \qquad \text{for} \quad i = 1, 2, \ldots, k-1, \tag{2.8}$$
$$\varphi_j = \eta_j \varphi_{j+1} + \mu_i \varphi_1 + p_{a_1} t \nu_j \varphi^1 + \xi_j \varphi, \qquad \text{for} \quad j = 1, 2, \ldots, l-1. \tag{2.9}$$

From (2.8) and (2.9), we have

$$\begin{aligned} \varphi^1 = {} & \left[\prod_{i=1}^{k-1} \alpha_i\right] x_0 + \left[\sum_{i=1}^{k-1} \left(\prod_{j=1}^{i-1} \alpha_j\right) \beta_i\right] \varphi^1 \\ & + \left[\sum_{i=1}^{k-1} \left(\prod_{j=1}^{i-1} \alpha_j\right) \gamma_i\right] p_{b_1} t \varphi_1 + \left[\sum_{i=1}^{k-1} \left(\prod_{j=1}^{i-1} \alpha_j\right) \delta_i\right] \varphi, \end{aligned} \tag{2.10}$$

$$\begin{aligned} \varphi_1 = {} & \left[\prod_{i=1}^{k-1} \eta_i\right] x_1 + \left[\sum_{i=1}^{k-1} \left(\prod_{j=1}^{i-1} \eta_j\right) \mu_i\right] \varphi_1 \\ & + \left[\sum_{i=1}^{k-1} \left(\prod_{j=1}^{i-1} \eta_j\right) \nu_i\right] p_{a_1} t \varphi^1 + \left[\sum_{i=1}^{k-1} \left(\prod_{j=1}^{i-1} \eta_j\right) \xi_i\right] \varphi. \end{aligned} \tag{2.11}$$

From (2.5), we obtain

$$\{1 - (1 - p_{a_1} - p_{b_1}) t\} \varphi = p_{a_1} t \varphi^1 + p_{b_1} t \varphi_1. \tag{2.12}$$

From (2.10) and (2.12), we have

$$\begin{aligned}
\varphi^1 &= \left[\prod_{i=1}^{k-1}\alpha_i\right]x_0 + \left[\sum_{i=1}^{k-1}\left(\prod_{j=1}^{i-1}\alpha_j\right)\beta_i\right]\varphi^1 \\
&\quad + \left[\sum_{i=1}^{k-1}\left(\prod_{j=1}^{i-1}\alpha_j\right)\gamma_i\right]\left[\{1-(1-p_{a_1}-p_{b_1})t\}\varphi - p_{a_1}t\varphi^1\right] \\
&\quad + \left[\sum_{i=1}^{k-1}\left(\prod_{j=1}^{i-1}\alpha_j\right)\delta_i\right]\varphi \\
&= \left[\prod_{i=1}^{k-1}\alpha_i\right]x_0 - \left[\sum_{i=1}^{k-1}\left(\prod_{j=1}^{i-1}\alpha_j\right)(p_{a_1}t\delta_i - \beta_i)\right]\varphi^1 \\
&\quad + \left[1 - \left(\prod_{i=1}^{k-1}\alpha_i\right) + \sum_{i=1}^{k-1}\left(\prod_{j=1}^{i-1}\alpha_j\right)(p_{a_1}t\delta_i - \beta_i)\right]\varphi.
\end{aligned}$$

From (2.11) and (2.12), we have

$$\begin{aligned}
\varphi_1 &= \left[\prod_{i=1}^{k-1}\eta_i\right]x_1 + \left[\sum_{i=1}^{k-1}\left(\prod_{j=1}^{i-1}\eta_j\right)\mu_i\right]\varphi_1 \\
&\quad + \left[\sum_{i=1}^{k-1}\left(\prod_{j=1}^{i-1}\eta_j\right)\nu_i\right]\left[\{1-(1-p_{a_1}-p_{b_1})t\}\varphi - p_{b_1}t\varphi_1\right] \\
&\quad + \left[\sum_{i=1}^{k-1}\left(\prod_{j=1}^{i-1}\eta_j\right)\xi_i\right]\varphi \\
&= \left[\prod_{i=1}^{k-1}\eta_i\right]x_1 - \left[\sum_{i=1}^{k-1}\left(\prod_{j=1}^{i-1}\eta_j\right)(p_{b_1}t\nu_i - \mu_i)\right]\varphi_1 \\
&\quad + \left[1 - \left(\prod_{i=1}^{k-1}\eta_i\right) + \sum_{i=1}^{k-1}\left(\prod_{j=1}^{i-1}\eta_j\right)(p_{b_1}t\nu_i - \mu_i)\right]\varphi.
\end{aligned}$$

Here, we set that

$$\begin{aligned}
H_0(t) &= 1 + \sum_{i=1}^{k-1}\left(\prod_{j=1}^{i-1}\alpha_j\right)(p_{a_1}t\gamma_i - \beta_i), \\
H_1(t) &= 1 + \sum_{i=1}^{l-1}\left(\prod_{j=1}^{i-1}\eta_j\right)(p_{b_1}t\nu_i - \mu_i).
\end{aligned}$$

Then, we have

$$\varphi^1 = \frac{(\prod_{i=1}^{k-1}\alpha_i)}{H_0(t)}x_0 + \left[1 - \frac{(\prod_{i=1}^{k-1}\alpha_i)}{H_0(t)}\right]\varphi, \tag{2.13}$$

$$\varphi_1 = \frac{(\prod_{i=1}^{l-1}\eta_i)}{H_1(t)}x_1 + \left[1 - \frac{(\prod_{i=1}^{l-1}\eta_i)}{H_1(t)}\right]\varphi. \tag{2.14}$$

From (2.5), (2.13) and (2.14), we obtain

$$\begin{aligned}\varphi = {} & \frac{p_{a_1}t(\prod_{i=1}^{k-1}\alpha_i)}{H_0(t)}x_0 + \frac{p_{b_1}t(\prod_{i=1}^{l-1}\eta_i)}{H_1(t)}x_1 \\ & + \left[t - \frac{p_{a_1}t(\prod_{i=1}^{k-1}\alpha_i)}{H_0(t)} - \frac{p_{b_1}t(\prod_{i=1}^{l-1}\eta_i)}{H_1(t)}\right]\varphi.\end{aligned} \tag{2.15}$$

We set

$$F_0(t) = \frac{p_{a_1}t(\prod_{i=1}^{k-1}\alpha_i)}{1 + \sum_{i=1}^{k-1}\frac{p_{a_1}t - A_it}{1 - B_it}(\prod_{j=1}^{i-1}\alpha_j)},$$

$$F_1(t) = \frac{p_{b_1}t(\prod_{i=1}^{l-1}\eta_i)}{1 + \sum_{i=1}^{l-1}\frac{p_{b_1}t - C_it}{1 - D_it}(\prod_{j=1}^{i-1}\eta_j)}.$$

From (2.15), we obtain

PROPOSITION 2.2. *The g.p.g.f. $\varphi(t)$ is given by*

$$\varphi(t) = \frac{F_0(t)x_0 + F_1(t)x_1}{1 - t + F_0(t) + F_1(t)}.$$

Example 2.3. If P_0 = “135” and P_1 = “246” in Proposition 2.2, we have

$$F_0(t) = p_1tp_3tp_5t,$$
$$F_1(t) = p_2tp_4tp_6t.$$

Then we obtain

$$\varphi(t) = \frac{p_1tp_3tp_5tx_0 + p_2tp_4tp_6tx_1}{1 - t + p_1tp_3tp_5t + p_2tp_4tp_6t}.$$

2.3 *Later waiting time problem for two patterns*

In this subsection, we consider the p.g.f. of the distribution of the waiting time for the later event between E_0 and E_1. Let τ be the waiting time for the later occurring event between E_0 and E_1. Let $\phi(t)$ be the p.g.f. of the distribution of the waiting time τ. We set

$$\Lambda(t) = \frac{p_{b_1}t(\prod_{i=1}^{l-1}\eta_i)}{1 + \sum_{i=1}^{l-1}\frac{p_{b_1}t - C_it}{1 - D_it}(\prod_{j=1}^{i-1}\eta_j)}.$$

Then, we have

PROPOSITION 2.4. *The p.g.f.* $\phi(t)$ *is given by*

$$\phi(t) = \frac{F_0(t)\dfrac{\Lambda(t)}{1-t+\Lambda(t)} + F_1(t)\dfrac{\Gamma(t)}{1-t+\Gamma(t)}}{1-t+F_0(t)+F_1(t)}$$

PROOF. In Proposition 2.2, x_0 can be regarded as a marker which means that the event E_0 occurs sooner. From the marker, we can see that the later occurring event is E_1. Here, we note that x_i is the p.g.f. of the conditional distribution of $\tau - t_0$ given that the sooner event E_i has just occurred at the t_0-th trial. In Lemma 2.1, we denote by $\varphi(t) = \varphi(t; \text{“}a_1 \cdots a_k\text{”})$ the p.g.f. of the distribution of the waiting time for the occurrence of the event E_0. Then, we have

$$x_0 = \varphi(t; \text{“}b_1 \cdots b_l\text{”}).$$

Similarly, we have

$$x_1 = \varphi(t; \text{“}a_1 \cdots a_k\text{”}).$$

This completes the proof.

3. Waiting time problems for countably many patterns

3.1 *The first occurrence problem for countably many patterns*

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, which take values in a countable set $S = \{0, 1, 2, \ldots\}$. The sequence $\{k_i\}_{i=0}^{\infty}$ is a given sequence of positive integers. Let $P_i = \text{“}a_{i,1}a_{i,2}\cdots a_{i,k_i}\text{”}$ for $i = 0, 1, 2, \ldots$ be the pattern of length k_i. Let $p_{i,j} = P\{X_1 = a_{i,j}\}$, $i = 0, 1, 2, \ldots$; $j = 1, 2, \ldots, k_i$. We assume that $\{P_i\} \bigcap \{P_j\} = \phi$ for $i \neq j$, and $a_{l,1} \le a_{l,2} \le \cdots \le a_{l,k_l}$ for $l = 0, 1, 2, \ldots$. For $i = 0, 1, 2, \ldots$, we denote by E_i the event that the pattern P_i occurs in the sequence $X_1, X_2, \ldots$. In this subsection, we consider the distribution of the waiting time for the first occurrence of an event of $\{E_i\}_{i=0}^{\infty}$. We derive a generalized probability generating function (g.p.g.f.) by adding markers x_i, $i = 0, 1, 2, \ldots$. Here, for each i, x_i represents that the first occurring event among $\{E_i\}_{i=0}^{\infty}$ is E_i. Let τ be the waiting time for the first occurrence of an event of $\{E_i\}_{i=0}^{\infty}$. Let $\phi(t)$ be the g.p.g.f. of the distribution of the waiting time τ. Let t_0 be an arbitrary positive integer and let $\phi_{i,j}(t)$ be the g.p.g.f. of the conditional distribution of $\tau - t_0$ given that the pattern $\text{“}a_{i,1}a_{i,2}\cdots a_{i,j}\text{”}$ has just occurred at the t_0-th trials, where $i = 0, 1, 2, \ldots$; $j = 1, 2, \ldots, k_i$ and $\phi_{i,k_i}(t) = x_i$. We also note that $\phi_{i,j}(t)$ does not depend on t_0. Here, we set that for $i = 0, 1, 2, \ldots$; $j = 1, 2, \ldots, k_i - 1$,

$$\begin{aligned} A_{i,j} &= p_{a_{i,1}} 1\{a_{i,j+1} \neq a_{i,1}\} 1\{a_{i,j} \neq a_{i,1}\}, \\ B_{i,j} &= p_{a_{i,1}} 1\{a_{i,j+1} \neq a_{i,1}\} 1\{a_{i,j} = a_{i,1}\}. \end{aligned}$$

Then ϕ and $\phi_{i,j}$, $i = 0, 1, 2, \ldots$; $j = 1, 2, \ldots, k_i$ satisfy the following system of linear equations:

$$\phi = \sum_{i=0}^{\infty} p_{i,1} t \phi_{i,1} + \left(1 - \sum_{i=0}^{\infty} p_{i,1}\right) t\phi, \tag{3.1}$$

$$\phi_{i,j} = p_{i,j+1}t\phi_{i,j+1} + \sum_{l=0,l\neq i}^{\infty} p_{l,1}t\phi_{l,1} + A_{i,j}t\phi_{i,1} + B_{i,j}t\phi_{i,j} + \left(1 - p_{i,j+1} - \sum_{l=0,l\neq i}^{\infty} p_{l,1} - A_{i,j} - B_{i,j}\right)t\phi, \quad \text{for} \quad i = 0,1,2,\ldots;\ j = 1,2,\ldots,k_i - 2, \tag{3.2}$$

$$\phi_{i,k_i-1} = p_{i,k_i}tx_i + \sum_{l=0,l\neq i}^{\infty} p_{l,1}t\phi_{l,1} + A_{i,k_i-1}t\phi_{i,1} + B_{i,k_i-1}t\phi_{i,k_i-1} + \left(1 - p_{i,k_i} - \sum_{l=0,l\neq i}^{\infty} p_{l,1} - A_{i,k_i-1} - B_{i,k_i-1}\right)t\phi, \quad \text{for} \quad i = 0,1,2,\ldots. \tag{3.3}$$

We set that for $i = 0,1,2,\ldots;\ j = 1,2,\ldots,k_i - 1$,

$$\alpha_{i,j} = \frac{p_{a_{i,j+1}}t}{1 - B_{i,j}t},$$

$$\beta_{i,j} = \frac{A_{i,j}t}{1 - B_{i,j}t},$$

$$\gamma_{i,j} = \frac{1}{1 - B_{i,j}t},$$

$$\delta_{i,j} = \frac{\left(1 - p_{a_{i,j+1}} - \sum_{\substack{l=0\\ l\neq i}}^{\infty} p_{l,1} - A_{i,j} - B_{i,j}\right)t}{1 - B_{i,j}t}.$$

From (3.2) and (3.3), we have

$$\phi_{i,j} = \alpha_{i,j}\phi_{i,j+1} + \beta_{i,j}\sum_{l=0,l\neq i}^{\infty} p_{l,1}t\phi_{l,1} + \gamma_{i,j}\phi_{i,1} + \delta_{i,j}\phi, \quad \text{for} \quad i = 0,1,2,\ldots;j = 1,2,\ldots,k_i - 2, \tag{3.4}$$

$$\phi_{i,k_i-1} = \alpha_{i,k_i-1}x_i + \beta_{i,k_i-1}\sum_{l=0,l\neq i}^{\infty} p_{l,1}t\phi_{l,1} + \gamma_{i,k_i-1}\phi_{i,1} + \delta_{i,k_i-1}\phi, \quad \text{for} \quad i = 0,1,2,\ldots. \tag{3.5}$$

From (3.4) and (3.5), we have, for each $i = 0,1,2,\ldots$,

$$\phi_{i,1} = \left[\prod_{m=1}^{k_i-1}\alpha_{i,m}\right]x_i + \left[\sum_{r=1}^{k_i-1}\left(\prod_{m=1}^{r-1}\alpha_{i,m}\right)\beta_{i,r}\right]\left[\sum_{l=0,l\neq i}^{\infty} p_{l,1}t\phi_{l,1}\right] + \sum_{r=1}^{k_i-1}\left(\prod_{m=1}^{r-1}\alpha_{i,m}\right)\gamma_{i,r}\phi_{i,1} + \sum_{r=1}^{k_i-1}\left(\prod_{m=1}^{r-1}\alpha_{i,r}\right)\delta_{i,r}\phi. \tag{3.6}$$

From (3.1), we have

$$\sum_{l=0,l\neq i}^{\infty} p_{l,1}\phi_{l,1} = \left[1-\left(1-\sum_{l=0}^{\infty}p_{l,1}\right)t\right]\phi - p_{i,1}t\phi_{i,1}. \tag{3.7}$$

From (3.6) and (3.7), we obtain, for each $i = 0,1,2,\dots,$

$$\begin{aligned}
\phi_{i,1} &= \left[\prod_{m=1}^{k_i-1}\alpha_{i,m}\right]x_i + \left[\sum_{r=1}^{k_i-1}\left(\prod_{m=1}^{r-1}\alpha_{i,m}\right)\beta_{i,r}\right]\left[1-\left(1-\sum_{l=0}^{\infty}p_{l,1}\right)t\right]\phi \\
&\quad - \left[\sum_{r=1}^{k_i-1}\left(\prod_{m=1}^{r-1}\alpha_{i,m}\right)\beta_{i,r}\right]p_{i,1}t\phi_{i,1} + \sum_{r=1}^{k_i-1}\left(\prod_{m=1}^{r-1}\alpha_{i,m}\right)\gamma_{i,r}\phi_{i,1} \\
&\quad + \sum_{r=1}^{k_i-1}\left(\prod_{m=1}^{r-1}\alpha_{i,r}\right)\delta_{i,r}\phi \\
&= \left[\prod_{m=1}^{k_i-1}\alpha_{i,m}\right]x_i - \sum_{r=1}^{k_i-1}\left(\prod_{m=1}^{r-1}\alpha_{i,m}\right)\frac{p_{i,1}t - A_{i,r}t}{1-B_{i,r}t}\phi_{i,1} \\
&\quad + \left[1-\left(\prod_{m=1}^{k_i-1}\alpha_{i,m}\right) + \sum_{r=1}^{k_i-1}\left(\prod_{m=1}^{r-1}\alpha_{i,m}\right)\frac{p_{i,1}t - A_{i,r}t}{1-B_{i,r}t}\right]\phi.
\end{aligned}$$

Here, we set

$$H_i(t) = 1 + \sum_{r=1}^{k_i-1}\left(\prod_{m=1}^{r-1}\alpha_{i,m}\right)\frac{p_{i,1}t - A_{i,r}t}{1-B_{i,r}t}, \qquad \text{for} \quad i = 0,1,2,\dots.$$

Then we have, for each $i = 0,1,2,\dots,$

$$\phi_{i,1} = \frac{[\prod_{m=1}^{k_i-1}\alpha_{i,m}]}{H_i(t)}x_i + \left\{1 - \frac{[\prod_{m=1}^{k_i-1}\alpha_{i,m}]}{H_i(t)}\right\}\phi. \tag{3.8}$$

From (3.1) and (3.8), we obtain

$$\begin{aligned}
\left[1-\left(1-\sum_{i=0}^{\infty}p_{i,1}\right)t\right]\cdot\phi &= \sum_{i=0}^{\infty}\frac{p_{i,1}t[\prod_{m=1}^{k_i-1}\alpha_{i,m}]}{H_i(t)}x_i + \sum_{i=0}^{\infty}p_{i,1}t\phi \\
&\quad + \sum_{i=0}^{\infty}\frac{p_{i,1}t[\prod_{m=1}^{k_i-1}\alpha_{i,m}]}{H_i(t)}\phi.
\end{aligned}$$

We set

$$G_i(t) = \frac{p_{i,1}t[\prod_{m=1}^{k_i-1}\alpha_{i,m}]}{H_i(t)}, \qquad \text{for} \quad i = 0,1,2,\dots.$$

Consequently, we have

THEOREM 3.1. *The g.p.g.f. $\phi(t)$ is given by*

$$\phi(t) = \frac{\sum_{i=0}^{\infty}G_i(t)x_i}{1-t+\sum_{i=0}^{\infty}G_i(t)}.$$

3.2 *The second occurrence problem for countably many patterns*

In this subsection, we consider the g.p.g.f. of the distribution of the waiting time for the occurrence of the second event among $\{E_i\}_{i=0}^{\infty}$. For each i and j satisfying $i \neq j$, $x_{i,j}$ denotes the marker which means that the first occurring event is E_i and the second occurring event is E_j among $\{E_i\}_{i=0}^{\infty}$. Let τ be the waiting time for the occurrence of the second event among $\{E_i\}_{i=0}^{\infty}$. Let $\phi(t)$ be the g.p.g.f. with markers $\{x_{i,j}\}$ of the distribution of the waiting time τ.

THEOREM 3.2. *The g.p.g.f. $\phi(t)$ is given by*

$$\phi(t) = \frac{\sum_{i=0}^{\infty} G_i(t) \frac{\sum_{\substack{j=0 \\ j\neq i}}^{\infty} G_j(t) x_{i,j}}{1 - t + \sum_{\substack{j=0 \\ j\neq i}}^{\infty} G_j(t)}}{1 - t + \sum_{i=0}^{\infty} G_i(t)}.$$

PROOF. In Theorem 3.1, x_i can be regarded as a marker which means that the first occurring event among $\{E_i\}_{i=0}^{\infty}$ is E_i. From the marker, we can see that the second occurring event is another event excepting the first event E_i among the events $\{E_i\}_{i=0}^{\infty}$. Here, we note that x_i is the p.g.f. of the conditional distribution of the waiting time $\tau - t_0$ given that the first event E_i has just occurred at the t_0-th trial. So we denote by x_i the g.p.g.f. of the distribution of the waiting time for the occurrence of the first event among $\{E_j\}_{j=0, j\neq i}^{\infty}$ From Theorem 3.1, we have, for each $i = 0, 1, 2, \ldots$,

$$x_i = \frac{\sum_{\substack{j=0 \\ j\neq i}}^{\infty} G_j(t) x_{i,j}}{1 - t + \sum_{\substack{j=0 \\ j\neq i}}^{\infty} G_j(t)}.$$

This completes the proof.

3.3 *The r-th occurrence problem for countably many patterns*

In this subsection, for each integer $r \geq 2$, we consider the g.p.g.f. of the distribution of the waiting time for the occurrence of the r-th event among $\{E_i\}_{i=0}^{\infty}$. We use markers $x_i, x_{i,j}, x_{i,j,l}, \ldots$ as in the previous subsections; for example, the marker $x_{i,j,l}$ means that the first occurring event is E_i, the second event E_j and the third event is E_l. We denote by $\phi_1 = \phi_1(t; x_{i_1}; i_1 = 0, 1, 2, \ldots)$ the g.p.g.f. of the distribution of the waiting time for the first occurrence among $\{E_i\}_{i=0}^{\infty}$. In general, $\phi_r = \phi_r(t; x_{i_1,i_2,\ldots,i_r}; i_1, i_2, \ldots, i_r = 0, 1, 2, \ldots)$ denotes the g.p.g.f. of the distribution of the waiting time τ for the occurrence of the r-th event among $\{E_i\}_{i=0}^{\infty}$. From the meaning of our problem, if $j \neq l$, then $i_j \neq i_l$ holds. We set

$$\psi_{i_1,i_2,\ldots,i_{r-1}}(t; x_{i_1,i_2,\ldots,i_{r-1},j}, j \neq i_1, i_2, \ldots, i_{r-1}) \equiv \frac{\sum_{\substack{j=0 \\ j\neq i_1,i_2,\ldots,i_{r-1}}}^{\infty} G_j(t) x_{i_1,i_2,\ldots,i_{r-1},j}}{1 - t + \sum_{\substack{j=0 \\ j\neq i_1,i_2,\ldots,i_{r-1}}}^{\infty} G_j(t)}.$$

Then, we have

THEOREM 3.3. *For each integer $r \geq 2$, it holds that*

$$\begin{aligned}&\phi_r(t; x_{i_1,i_2,\ldots,i_r}, i_1, \ldots, i_r = 0, 1, 2, \ldots)\\&= \phi_{r-1}(t; x_{i_1,i_2,\ldots,i_{r-1}} = \psi_{i_1,i_2,\ldots,i_{r-1}}(t; x_{i_1,i_2,\ldots,i_{r-1},j})).\end{aligned}$$

The right hand side of this equation means the formula which is obtained by replacing every marker $x_{i_1,i_2,\ldots,i_{r-1}}$ in $\phi_{r-1}(t; x_{i_1,i_2,\ldots,i_{r-1}})$ by $\psi_{i_1,i_2,\ldots,i_{r-1}}(t; x_{i_1,i_2,\ldots,i_{r-1},j})$.

PROOF. In Theorem 3.2, $x_{i,j}$ can be regarded as a marker which means that the first occurring event is E_i and the second occurring event is E_j among $\{E_i\}_{i=0}^{\infty}$. Similarly, $x_{i_1,i_2,\ldots,i_{r-1}}$ can be regarded as a marker which means that the l-th occurring event is E_{i_l} for $l = 1, 2, \ldots, r-1$ among $\{E_i\}_{i=0}^{\infty}$. From the marker, we can see that the r-th occurring event is another event excepting the events $E_{i_1}, E_{i_2}, \ldots, E_{i_{r-1}}$ among $\{E_i\}_{i=0}^{\infty}$. Here, we note that $x_{i_1,i_2,\ldots,i_{r-1}}$ is the p.g.f. of the conditional distribution of the waiting time $\tau - t_0$ given that the events $E_{i_1}, E_{i_2}, \ldots, E_{i_{r-2}}$ have already occurred until t_0-th trial and the event $E_{i_{r-1}}$ has just occurred at the t_0-th trial. So we denote by $x_{i_1,i_2,\ldots,i_{r-1}}$ the g.p.g.f. of the distribution of the waiting time for the occurrence of the first event among $\{E_j\}_{j=0,j\neq i_1,i_2,\ldots,i_{r-1}}^{\infty}$. From Theorem 3.1, we have

$$x_{i_1,i_2,\ldots,i_{r-1}} = \frac{\sum_{\substack{j=0\\ j\neq i_1,i_2,\ldots,i_{r-1}}}^{\infty} G_j(t) x_{i_1,i_2,\ldots,i_{r-1},j}}{1 - t + \sum_{\substack{j=0\\ j\neq i_1,i_2,\ldots,i_{r-1}}}^{\infty} G_j(t)}.$$

This completes the proof.

4. Number of occurrence of sub-pattern in higher order Markov chain

In this Section, let k, l and m be fixed positive integers such that $m \leq l < k$. Let $X_{-m+1}, X_{-m+2}, \ldots, X_0, X_1, \ldots$ be a $\{0, 1, 2, \ldots\}$-valued m-th order Markov chain with

$$\begin{aligned}\pi_{a_1,\ldots,a_m} &= P[X_{-m+1} = a_1, X_{-m+2} = a_2, \ldots, X_0 = a_m],\\ p_{a_{m+1}|a_1,\ldots,a_m} &= P[X_i = a_{m+1} | X_{i-m} = a_1, X_{i-m+1} = a_2, \ldots, X_{i-1} = a_m],\end{aligned}$$

for $a_1, \ldots, a_{m+1} = 0, 1, 2, \ldots$ and $i = 1, 2, \ldots$. For $a_1, \ldots, a_{m+1} = 0, 1, 2, \ldots$, we assume that $0 < p_{a_{m+1}|a_1,\ldots,a_m} < 1$. Let $P_0 =$ "$a_1 a_2 \cdots a_k$" be a pattern of length k, where $a_1 \leq a_2 \leq \cdots \leq a_k$. We denote by $S_l =$ "$a_1 a_2 \cdots a_l$" the sub-pattern of P_0 for $l = m, m+1, \ldots, k-1$. Let E_0 be the event that the pattern P_0 occurs in the sequence $X_1, X_2, \ldots$. We denote by τ the waiting time for the occurrence of the event E_0. We derive the distributions of the number of occurrences of sub-patterns of length l, that is, S_l until τ and if the sub-pattern S_l is a run, we obtain the distributions of the number of overlapping occurrences of runs of length l until τ. Let $\phi^{a_1,\ldots,a_m}(t)$ be the probability generating function (p.g.f.) of the conditional distribution of the number of occurrences of sub-patterns of length l until τ given that $X_{-m+1} = a_1, X_{-m+2} = a_2, \ldots, X_0 = a_m$. Suppose we have currently sub-pattern of length l, that is, "$a_1 a_2 \cdots a_l$" in $X_i, X_{i-1}, \ldots$. Then, we denote by $\phi_i(t)$ the p.g.f. of the conditional distribution of the number of occurrences of sub-patterns of length l from this time until τ. We are waiting for the first occurrence of the pattern P_0 of length k. Then, starting from any initial state $(a_1, \ldots, a_m)$ we observe the first occurrence of the sub-pattern S_l of length l somewhere in $X_l, X_{l+1}, \ldots$ with probability 1. By considering the m-th order Markov chain just after the first occurrence of the sub-pattern S_l, we see that

$$\phi^{(a_1,\ldots,a_m)} = t\phi_l \qquad \text{for each initial state } (a_1, \ldots, a_m).$$

This shows that all the conditional distributions are equal to each other and they do not depend on their initial conditions. Hence, we denote $\phi \equiv t\phi_l = \phi^{(a_1,\ldots,a_m)}$. Then it is easy to see that

$$\begin{cases} \phi_l = p_{a_{l+1}|a_{l-m+1}\cdots a_l} A_{l+1}(t)\phi_{l+1} + p_{a_l|a_{l-m+1}\cdots a_l} B_{l+1} t\phi_l \\ \qquad + (1 - p_{a_{l+1}|a_{l-m+1}\cdots a_l} - p_{a_l|a_{l-m+1}\cdots a_l} B_{l+1})\phi, \\ \phi_{l+1} = p_{a_{l+2}|a_{l-m+2}\cdots a_{l+1}} A_{l+2}(t)\phi_{l+2} + p_{a_{l+1}|a_{l-m+2}\cdots a_{l+1}} B_{l+2} t\phi_{l+1} \\ \qquad + (1 - p_{a_{l+2}|a_{l-m+2}\cdots a_{l+1}} - p_{a_{l+1}|a_{l-m+2}\cdots a_{l+1}} B_{l+2})\phi, \\ \quad \vdots \\ \phi_{k-1} = p_{a_k|a_{k-m}\cdots a_{k-1}} A_k(t) + p_{a_{k-1}|a_{k-m}\cdots a_{k-1}} B_k t\phi_{k-1} \\ \qquad + (1 - p_{a_k|a_{k-m}\cdots a_{k-1}} - p_{a_{k-1}|a_{k-m}\cdots a_{k-1}} B_k)\phi, \end{cases}$$

where for $i = l+1, \ldots, k$,

$$A_i(t) = 1\{a_i = a_1\}t + 1\{a_i \neq a_1\},$$
$$B_i = 1\{a_{i-1} = a_1\}1\{a_i \neq a_1\}.$$

Here, we set that for $i = l+1, \ldots, k$,

$$\alpha_i(t) = \frac{p_{a_i|a_{i-m}\cdots a_{i-1}} A_i(t)}{1 - p_{a_{i-1}|a_{i-m}\cdots a_{i-1}} B_i t},$$
$$\beta_i(t) = \frac{1 - p_{a_i|a_{i-m}\cdots a_{i-1}} - p_{a_{i-1}|a_{i-m}\cdots a_{i-1}} B_i}{1 - p_{a_{i-1}|a_{i-m}\cdots a_{i-1}} B_i t}.$$

Then, we have the system of linear equations:

$$\begin{cases} \phi_l = \alpha_{l+1}(t)\phi_{l+1} + \beta_{l+1}(t)\phi, \\ \phi_{l+1} = \alpha_{l+2}(t)\phi_{l+2} + \beta_{l+2}(t)\phi, \\ \quad \vdots \\ \phi_{k-1} = \alpha_k(t) + \beta_k(t)\phi. \end{cases}$$

By solving this system of linear equations, we obtain

$$\phi_l = \left[\prod_{i=l+1}^{k} \alpha_i(t)\right] + \sum_{i=l+1}^{k} \left[\prod_{j=l+1}^{i-1} \alpha_j(t)\right] \beta_i(t)\phi.$$

Then, $\phi = t\phi_l$ implies the following result.

THEOREM 4.1. *The p.g.f.* $\phi(t)$ *is given by*

$$\phi = \frac{[\prod_{i=l+1}^{k} \alpha_i(t)]t}{1 - t\sum_{i=l+1}^{k}[\prod_{j=l+1}^{i-1} \alpha_j(t)]\beta_i(t)},$$

where for $i = l+1, \ldots, k$,

$$\alpha_i(t) = \frac{p_{a_i|a_{i-m}\cdots a_{i-1}}(1\{a_i = a_1\}t + 1\{a_i \neq a_1\})}{1 - p_{a_{i-1}|a_{i-m}\cdots a_{i-1}}1\{a_{i-1} = a_1\}1\{a_i \neq a_1\}t},$$
$$\beta_i(t) = \frac{1 - p_{a_i|a_{i-m}\cdots a_{i-1}} - p_{a_{i-1}|a_{i-m}\cdots a_{i-1}}1\{a_{i-1} = a_1\}1\{a_i \neq a_1\}}{1 - p_{a_{i-1}|a_{i-m}\cdots a_{i-1}}1\{a_{i-1} = a_1\}1\{a_i \neq a_1\}t}.$$

COROLLARY 4.2. *If there exist some* i, j $(= 1, 2, \ldots, l)$ *such that* $a_i < a_j$ $(i < j)$, *then the distribution of the number of occurrences of the sub-patterns of length* l *until the first occurrence of the pattern of length* k *in* $X_1, X_2, \ldots$ *is the geometric distribution of order* 1, $G_1(\prod_{i=l+1}^{k} p_{a_i|a_{i-m}\cdots a_{i-1}})$.

PROOF. By setting $\alpha_i(t) = p_{a_i|a_{i-m}\cdots a_{i-1}}$, $\beta_i(t) = 1 - p_{a_i|a_{i-m}\cdots a_{i-1}}$ for $i = l+1, \ldots, k$ in Theorem 4.1, we obtain

$$\begin{aligned}\phi &= \frac{[\prod_{i=l+1}^{k} p_{a_i|a_{i-m}\cdots a_{i-1}}]t}{1 - t\sum_{i=l+1}^{k}[\prod_{j=l+1}^{i-1} p_{a_j|a_{j-m}\cdots a_{j-1}}](1 - p_{a_i|a_{i-m}\cdots a_{i-1}})} \\ &= \frac{[\prod_{i=l+1}^{k} p_{a_i|a_{i-m}\cdots a_{i-1}}]t}{1 - t + [\prod_{i=l+1}^{k} p_{a_i|a_{i-m}\cdots a_{i-1}}]t}.\end{aligned}$$

This completes the proof.

COROLLARY 4.3. *If* $a_1 = a_2 = \cdots = a_k$, *then the distribution of the number of overlapping occurrences of "*a_1*"-runs of length* l *until the first occurrence of "*a_1*"-run of length* k *in* $X_1, X_2, \ldots$ *is the shifted geometric distribution of order* $k-l$, $G_{k-l}(p_{a_1|a_1\cdots a_1}, k-l+1)$.

PROOF. By setting $\alpha_i(t) = p_{a_1|a_1\cdots a_1}t$, $\beta_i(t) = 1 - p_{a_1|a_1\cdots a_1}$ for $i = l+1, \ldots, k$ in Theorem 4.1, we obtain

$$\begin{aligned}\phi &= \frac{[\prod_{i=l+1}^{k} p_{a_1|a_1\cdots a_1}t]t}{1 - t\sum_{i=l+1}^{k}[\prod_{j=l+1}^{i-1} p_{a_1|a_1\cdots a_1}t](1 - p_{a_1|a_1\cdots a_1})} \\ &= \frac{p_{a_1|a_1\cdots a_1}^{k-l}t^{k-l+1}(1 - p_{a_1|a_1\cdots a_1}t)}{1 - t + (1 - p_{a_1|a_1\cdots a_1})t(p_{a_1|a_1\cdots a_1}t)^{k-l}}.\end{aligned}$$

This completes the proof.

Remark. If we assume that X's are $\{0, 1\}$-valued and let $a_1 = 1$, then Corollary 4.3 agrees with Theorem 23.2.1 of Hirano *et al.* (1997).

COROLLARY 4.4. *If $a_1 = a_2 = \cdots = a_l < a_{l+1} \leq \cdots \leq a_k$, then the p.g.f. of the distribution of the number of overlapping occurrences of "a_1"-runs of length l until the first occurrence of the pattern of length k in $X_1, X_2, \ldots$ is given by*

$$\phi(t) = \frac{[\prod_{i=l+1}^{k} p_{a_i|a_{i-m}\cdots a_{i-1}}]t^2}{(1-t)(1+p_{a_{l+1}|a_{l-m-1}\cdots a_l}t) + [\prod_{i=l+1}^{k} p_{a_i|a_{i-m}\cdots a_{i-1}}]t^2}.$$

PROOF. In Theorem 4.1, we set that

$$\alpha_{l+1}(t) = \frac{p_{a_{l+1}|a_{l-m+1}\cdots a_l}t}{1-p_{a_l|a_{l-m-1}\cdots a_l}t},$$
$$\beta_{l+1}(t) = \frac{1-p_{a_{l+1}|a_{l-m+1}\cdots a_l} - p_{a_l|a_{l-m+1}\cdots a_l}}{1-p_{a_l|a_{l-m-1}\cdots a_l}t},$$

and for $i = l+2, \ldots, k$,

$$\alpha_i(t) = p_{a_i|a_{i-m}\cdots a_{i-1}},$$
$$\beta_i(t) = 1 - p_{a_i|a_{i-m}\cdots a_{i-1}}.$$

From Theorem 4.1, we have

$$\begin{aligned}\phi &= \frac{\alpha_{l+1}(t)[\prod_{i=l+2}^{k}\alpha_i(t)]t}{1-t(\beta_{l+1}(t)+\sum_{i=l+2}^{k}[\prod_{j=l+1}^{i-1}\alpha_j(t)]\beta_i(t))} \\ &= \frac{\alpha_{l+1}(t)[\prod_{i=l+2}^{k}\alpha_i(t)]t}{1-t(\beta_{l+1}(t)+\alpha_{l+1}(t)\sum_{i=l+2}^{k}[\prod_{j=l+2}^{i-1}\alpha_j(t)]\beta_i(t))} \\ &= \frac{\alpha_{l+1}(t)[\prod_{i=l+2}^{k}\alpha_i(t)]t}{1-t(\beta_{l+1}(t)+\alpha_{l+1}(t)[1-\prod_{i=l+2}^{k}\alpha_i(t)])} \\ &= \frac{\alpha_{l+1}(t)[\prod_{i=l+2}^{k}\alpha_i(t)]t}{1-(\alpha_{l+1}(t)+\beta_{l+1}(t))t+\alpha_{l+1}(t)[\prod_{i=l+2}^{k}\alpha_i(t)]t} \\ &= \frac{\alpha_{l+1}(t)[\prod_{i=l+2}^{k}\alpha_i(t)]t}{\frac{(1-t)(1+p_{a_{l+1}|a_{l-m-1}\cdots a_l}t)}{1-p_{a_l|a_{l-m-1}\cdots a_l}t}+\alpha_{l+1}(t)[\prod_{i=l+2}^{k}\alpha_i(t)]t}.\end{aligned}$$

This completes the proof.

Acknowledgements

The author wishes to thank the referees and the editor for their valuable comments.

References

Aki, S. (1992). Waiting time problems for a sequence of discrete random variables, *Ann. Inst. Statist. Math.*, **44**, 363–378.

Aki, S. and Hirano, K. (1993). Discrete distributions related to succession events in a two-state Markov chain, *Statistical Sciences and Data Analysis*; *Proceedings of the Third Pacific Area Statistical Conference* (eds. K. Matusita, M. L. Puri and T. Hayakawa), 467–474, VSP International Science Publishers, Zeist.

Aki, S. and Hirano, K. (1994). Distributions of numbers of failures and successes until the first consecutive k successes, *Ann. Inst. Statist. Math.*, **46**, 193–202.

Aki, S. and Hirano, K. (1995). Joint distributions of numbers of success-runs and failures until the first consecutive k successes, *Ann. Inst. Statist. Math.*, **47**, 225–235.

Balasubramanian, K., Viveros, R. and Balakrishnan, N. (1993). Sooner and later waiting time problems for Markovian Bernoulli trials, *Statist. Probab. Lett.*, **18**, 153–161.

Ebneshahrashoob, M. and Sobel, M. (1990). Sooner and later waiting time problems for Bernoulli trials: Frequency and run quotas, *Statist. Probab. Lett.*, **9**, 5–11.

Fu, J. C. (1996). Distribution theory of runs and patterns associated with a sequence of multi-state trials, *Statistica Sinica*, **6**, 957–974.

Fu, J. C. and Koutras, M. V. (1994). Distribution theory of runs: A Markov chain approach, *J. Amer. Statist. Assoc.*, **89**, 1050–1058.

Gerber, H. U. and Li, S.-Y. R. (1981). The occurrence of sequence patterns in repeated experiments and hitting times in a Markov chain, *Stoch. Proc. Appl.*, **11**, 101–108.

Hahn, G. J. and Gage, J. B. (1983). Evaluation of a start-up demonstration test, *Journal of Quality Technology*, **15**, 103–106.

Hirano, K., Aki, S. and Uchida, M. (1996). 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* (ed. N. Balakrishnan), 401–410, Barikhäuser Boston.

Li, S.-Y. R. (1980). A martingale approach to the study of occurrence of sequence patterns in repeated experiments, *Ann. Probab.*, **8**, 1171–1176.

Ling, K, D. (1990). On geometric distributions of order $(k_1, \ldots, k_m)$, *Statist. Probab. Lett.*, **9**, 163–171.

Philippou, A. N., Georghiou, C. and Philippou, G. N. (1983). A generalized geometric distribution and some of its properties, *Statist. Probab. Lett.*, **1**, 171–175.

Uchida, M. and Aki, S. (1995). Sooner and later waiting time problems in a two-state Markov chain, *Ann. Inst. Statist. Math.*, **44**, 363–378.

Viveros, R. and Balakrishnan, N. (1993). Statistical inference from start-up demonstration test data, *Journal of Quality Technology*, **25**, 119–130.