Ann. Inst. Statist. Math.
Vol. 56, No. 1, 143–168 (2004)
©2004 The Institute of Statistical Mathematics

JOINT DISTRIBUTIONS ASSOCIATED WITH PATTERNS, SUCCESSES AND FAILURES IN A SEQUENCE OF MULTI-STATE TRIALS*

KIYOSHI INOUE

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

(Received May 29, 2002; revised November 7, 2002)

Abstract. Let $\{Z_t, t \geq 1\}$ be a sequence of trials taking values in a given set $\mathcal{A} = \{0, 1, 2, \dots, m\}$, where we regard the value 0 as failure and the remaining m values as successes. Let $\mathcal{E}$ be a (single or compound) pattern. In this paper, we provide a unified approach for the study of two joint distributions, i.e., the joint distribution of the number X_n of occurrences of $\mathcal{E}$, the numbers of successes and failures in n trials and the joint distribution of the waiting time T_r until the r-th occurrence of $\mathcal{E}$, the numbers of successes and failures appeared at that time. We also investigate some distributions as by-products of the two joint distributions. Our methodology is based on two types of the random variables X_n (a Markov chain imbeddable variable of binomial type and a Markov chain imbeddable variable of returnable type). The present work develops several variations of the Markov chain imbedding method and enables us to deal with the variety of applications in different fields. Finally, we discuss several practical examples of our results.

Key words and phrases: Run, pattern, waiting time, enumeration schemes, Markov chain, double generating function, probability generating function, Markov chain imbedding method, transition probability matrices.

1. Introduction

Let $\{Z_t, t \geq 1\}$ be a sequence of trials taking values in a given set $\mathcal{A} = \{0, 1, 2, \dots, m\}$. We regard the value 0 as failure and the remaining m values as successes. If $m = 1$ then the sequence can be regarded as the Bernoulli trials. Let $\mathcal{E}$ be a (single or compound) pattern whose elements are integers in $\mathcal{A}$. There are two important distributions associated with the pattern $\mathcal{E}$, which are applied to a wide range of areas (for example, quality control, reliability theory, psychology, genome sequence analysis, etc). One is the distribution of the number X_n of occurrences of the pattern $\mathcal{E}$ among $Z_1, Z_2, \dots, Z_n$. The other is the distribution of the waiting time T_r until the r-th ($r \geq 1$) occurrence of $\mathcal{E}$. The distributions of X_n and T_r have been studied by many authors in various situations (see, for example, Biggins and Cannings (1987), Blom and Thorburn (1982), Robin and Daudin (1999), Fu (1996) and Uchida (1998)).

Recently, Fu and Koutras (1994) introduced a finite Markov chain imbedding method for the study of run-related problems, which has a great potential for extending to other problems (see Koutras (1996a), Fu and Lou (2000) and Koutras and Alexandrou (1997a)). Koutras and Alexandrou (1995) refined this method and introduced a Markov

*This research was partially supported by the ISM Cooperative Research Program (2002-ISM·CRP-2007).

chain imbeddable variable of binomial type (M.V.B.). They studied X_n in the case where $\mathcal{E}$ is a success run (non-overlapping, at least, overlapping) of length k in a sequence of n Bernoulli trials. Han and Aki (1999) introduced a Markov chain imbeddable variable of returnable type (M.V.R.) and derived the distribution of the number of success runs of exact length k. By using the Markov chain imbedding method, Inoue and Aki (2003) studied T_r in the case where $\mathcal{E}$ is a ℓ-overlapping success run of length k in a sequence of Markov dependent trials.

Besides the two random variables X_n and T_r, other important random variables are the number of outcomes "i" ($i = 0, 1, \ldots, m$) in the observed sequence. Let $S_{n,i}$ be the number of the outcomes "i" ($i = 1, 2, \ldots, m$) and let $F_{n,0}$ be the number of the outcomes "0" among $Z_1, Z_2, \ldots, Z_n$. We can obtain useful information from the two joint distributions: the joint distribution of the number X_n of occurrences of the pattern $\mathcal{E}$, the numbers $S_{n,i}$ of outcomes "i" ($i = 1, 2, \ldots, m$) and the number $F_{n,0}$ of outcomes "0" among $Z_1, Z_2, \ldots, Z_n$ or the joint distribution of the waiting time T_r until r-th occurrence of the pattern $\mathcal{E}$, the numbers $S_{T_r,i}$ of outcomes "i" ($i = 1, 2, \ldots, m$) and the number $F_{T_r,0}$ of outcomes "0" among $Z_1, Z_2, \ldots, Z_{T_r}$. For example, in quality control, it is quite natural in sampling inspection to use a run of defective items as a stopping criterion (see Koutras (1997)). Each item is classified to three categories: fully conformable (type S^*), partially conformable (type F) and totally rejectable (type S). Assume that we decide to accept the lot if k_1 consecutive S^*-type items are observed and reject the lot if k_2 consecutive S-type items are observed. Then the distribution of the numbers of items of types S^*, S and F observed until the termination of the sampling inspection plan is used to take corrective action on the production line. For example, many problems in bioinformatics relate to the comparison of two (or more) DNA sequences taking values in $\mathcal{A} = \{A, C, G, T\}$. In order to compare two sequences, we should extract information from these sequences composed of four letters. If we consider the test of the hypothesis that the two sets of probabilities for the four letters are equal, we need to count the numbers of the four letters A, C, G and T, respectively in the observed sequences. The distribution of the frequencies of the four letters as well as the distribution of the number of occurrences of $\mathcal{E}$ give more insight into the analysis of the DNA sequence (see Ewens and Grant (2001)).

We should make extensive use of the Markov chain imbedding method, in order to obtain the joint distributions of $(X_n, F_{n,0}, S_{n,1}, \ldots, S_{n,m})$, $(T_r, F_{T_r,0}, S_{T_r,1}, \ldots, S_{T_r,m})$. The purpose in this paper is to develop a general workable framework for the Markov chain imbedding method for the derivation of the joint probability distribution functions and the joint probability generating functions (pgf's) of $(X_n, F_{n,0}, S_{n,1}, \ldots, S_{n,m})$, $(T_r, F_{T_r,0}, S_{T_r,1}, \ldots, S_{T_r,m})$ in the various ways of counting runs and patterns. We can deal with the wide class of patterns by using the results in this paper, since we consider a sequence of trials with more than two outcomes and compound patterns.

The present paper is organized as follows. In Section 2, we introduce necessary definitions and notations. In this paper, each one of the two cases (the variable X_n is an M.V.B., the variable X_n is an M.V.R.) is treated separately. In Section 3, in the case where the variable X_n is an M.V.B., we develop a unified approach for the study of the joint distribution of $(X_n, S_{n,1}, \ldots, S_{n,m})$. Through the joint distribution of $(X_n, S_{n,1}, \ldots, S_{n,m})$, we consider the joint distribution of $(X_n, F_{n,0}, S_{n,1}, \ldots, S_{n,m})$. As by-products, we examine the joint distributions of $(X_n, S_{n,i})$ ($i = 1, 2, \ldots, m$), $(X_n, F_{n,0})$ (and also, the marginal distribution of X_n). The formulae for the expected value of X_n are also obtained. In Section 4, in the case where the variable X_n is an M.V.B., we study the joint distribution of $(T_r, S_{T_r,1}, \dots, S_{T_r,m})$. Through the joint distribution of $(T_r, S_{T_r,1}, \dots, S_{T_r,m})$, we consider the joint distributions of $(T_r, F_{T_r,0}, S_{T_r,1}, \dots, S_{T_r,m})$. As by products, the joint distributions of $(T_r, S_{T_r,i})$ $(i = 1, 2, \dots, m)$, $(T_r, F_{T_r,0})$ (and also, the marginal distributions of T_r, $S_{T_r,i}$ $(i = 1, 2, \dots, m)$, $F_{T_r,0}$) are considered. The formulae for the expected value of T_r are also obtained. In Section 5, in the case where the variable X_n is an M.V.R., we study the joint distribution of $(X_n, S_{n,1}, \dots, S_{n,m})$. Through the joint distribution of $(X_n, S_{n,1}, \dots, S_{n,m})$, we consider the joint distributions of $(X_n, F_{n,0}, S_{n,1}, \dots, S_{n,m})$. As by-products, we examine the joint distributions of $(X_n, S_{n,i})$ $(i = 1, 2, \dots, m)$, $(X_n, F_{n,0})$ (and also, the marginal distribution of X_n). The formulae for the expected value of X_n are also obtained. Finally, in Section 6, we discussed several practical applications.

2. Definitions and notations

Let $\{Z_t, t \geq 1\}$ be a sequence of trials taking values in a given set $\mathcal{A} = \{0, 1, 2, \dots, m\}$, where we regard the value 0 as failure and the remaining m values as successes. Let $\mathcal{E}$ be any pattern (simple or compound) whose elements are integers in $\mathcal{A}$ and let n_0 be the number of 0 element which the pattern $\mathcal{E}$ contains. In the sequel, we assume that the length of pattern $\mathcal{E}$ is greater than 1. In practice, this is the most common situation. Then, we denote the number of occurrences of $\mathcal{E}$ by X_n and denote the number of outcomes "i" $(i = 1, 2, \dots, m)$ by $S_{n,i}$ among $Z_1, \dots, Z_n$ (n a fixed integer). We denote the joint probability distribution function of X_n and $\boldsymbol{S}_n (= (S_{n,1}, S_{n,2}, \dots, S_{n,m}))$ by

$$\begin{aligned} f_n(x, \boldsymbol{y}) &= \Pr(X_n = x, S_{n,1} = y_1, S_{n,2} = y_2, \dots, S_{n,m} = y_m), \\ &= \Pr(X_n = x, \boldsymbol{S}_n = \boldsymbol{y}), \qquad x = 0, 1, \dots, \ell_n \quad \text{and} \quad 0 \leq y_1, \dots, y_m \leq n - xn_0, \end{aligned}$$

where, ℓ_n is the maximum number of occurrences of $\mathcal{E}$ that can be accommodated in n trials, that is, $\ell_n = \max\{x : \Pr(X_n = x) > 0\}$. Needless to say, under the assumption that the length of pattern $\mathcal{E}$ is greater than 1, we have $\ell_1 = 0$ and $f_1(x, \boldsymbol{y}) = 0$ for $x \neq 0$.

The corresponding joint pgf and double generating function of $(X_n, S_{n,1}, \dots, S_{n,m})$ will be defined by

$$\phi_n(u, \boldsymbol{v}) = E(u^{X_n} v_1^{S_{n,1}} \cdots v_m^{S_{n,m}}) = \sum_{x=0}^{\ell_n} \sum_{\boldsymbol{y}=\boldsymbol{0}}^{n-xn_0} f_n(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}}, \tag{2.1}$$

$$\Phi(u, \boldsymbol{v}; w) = \sum_{n=1}^{\infty} \phi_n(u, \boldsymbol{v}) w^n = \sum_{n=1}^{\infty} \sum_{x=0}^{\ell_n} \sum_{\boldsymbol{y}=\boldsymbol{0}}^{n-xn_0} f_n(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} w^n, \tag{2.2}$$

respectively, where $\sum_{\boldsymbol{y}=\boldsymbol{0}}^{n-xn_0} = \sum_{y_1=0}^{n-xn_0} \cdots \sum_{y_m=0}^{n-xn_0}$ and $\boldsymbol{v}^{\boldsymbol{y}} = v_1^{y_1} \cdots v_m^{y_m}$.

Let us denote by T_r, $(r \geq 1)$ the waiting time for the r-th occurrence of $\mathcal{E}$ and its joint probability distribution function of $(T_r, S_{T_r,1}, \dots, S_{T_r,m})$ by

$$h_r(n, \boldsymbol{y}) = \Pr(T_r = n, \boldsymbol{S}_{T_r} = \boldsymbol{y}), \qquad n = 1, 2, \dots \quad \text{and} \quad 0 \leq y_1, \dots, y_m \leq n - rn_0.$$

The corresponding joint pgf and double generating function of $(T_r, S_{T_r,1}, \dots, S_{T_r,m})$ will be defined by

$$H_r(w, \boldsymbol{v}) = E(w^{T_r} v_1^{S_{T_r,1}} \cdots v_m^{S_{T_r,m}}) = \sum_{n=1}^{\infty} \sum_{\boldsymbol{y}=\boldsymbol{0}}^{n-rn_0} h_r(n, \boldsymbol{y}) w^n \boldsymbol{v}^{\boldsymbol{y}}, \tag{2.3}$$

$$H(w, \boldsymbol{v}; u) = \sum_{r=1}^{\infty} H_r(w, \boldsymbol{v})u^r = \sum_{r=1}^{\infty}\sum_{n=1}^{\infty}\sum_{\boldsymbol{y}=\boldsymbol{0}}^{n-rn_0} h_r(n, \boldsymbol{y})w^n \boldsymbol{v}^{\boldsymbol{y}} u^r, \tag{2.4}$$

respectively.

Let X_n be a non-negative integer valued random variable taking on the values $0, 1, \ldots, \ell_n$. Then, according to Koutras and Alexandrou (1995), the random variable X_n is called a Markov chain imbeddable variable of binomial type (M.V.B.) if

(1) There exists a Markov chain $\{Y_t, t \geq 0\}$ defined on a state space Ω.

(2) There exists a partition $\{\boldsymbol{U}_x : x \geq 0\}$ on the state space Ω.

(3) For every x, $P(X_n = x) = P(Y_n \in \boldsymbol{U}_x)$.

(4) $P(Y_t \in \boldsymbol{U}_w \mid Y_{t-1} \in \boldsymbol{U}_x) = 0$ if $w \neq x, x+1$ and $t \geq 1$.

Assume first that the sets $\boldsymbol{U}_x$ of the partition $\{\boldsymbol{U}_x, x \geq 0\}$ have the same cardinality $s = |\boldsymbol{U}_x|$ for every x, more specifically $\boldsymbol{U}_x = \{U_{x,0}, U_{x,1}, \ldots, U_{x,s-1}\}$. According to Han and Aki (1999), the random variable X_n is called a Markov chain imbeddable variable of returnable type (M.V.R.), if all of the conditions in the above definition of M.V.B. hold, except that (4) is replaced by the following statement:

(4′) $P(Y_t \in \boldsymbol{U}_w \mid Y_{t-1} \in \boldsymbol{U}_x) = 0$ if $w \neq x-1, x, x+1$ and $t \geq 1$.

We denote the initial probabilities of the Markov chain by

$$\boldsymbol{\pi}_x = (\Pr(Y_0 = U_{x,0}), \Pr(Y_0 = U_{x,1}), \ldots, \Pr(Y_0 = U_{x,s-1})), \qquad x \geq 0,$$

and denote the probability vectors

$$\boldsymbol{f}_t(x) = (\Pr(Y_t = U_{x,0}), \Pr(Y_t = U_{x,1}), \ldots, \Pr(Y_t = U_{x,s-1})), \qquad 0 \leq x \leq \ell_n,\ 1 \leq t \leq n.$$

Clearly, we have

$$\Pr(X_n = x) = \boldsymbol{f}_n(x)\boldsymbol{1}', \qquad 0 \leq x \leq \ell_n,$$

where, $\boldsymbol{1} = (1, 1, \ldots, 1) \in \mathcal{R}^s$.

We introduce the next two matrices,

(i) the within states matrix $A_t(x) = (\Pr(Y_t = U_{x,j} \mid Y_{t-1} = U_{x,i}))_{s\times s}$,

(ii) the between states matrix $B_t(x) = (\Pr(Y_t = U_{x+1,j} \mid Y_{t-1} = U_{x,i}))_{s\times s}$.

If X_n is an M.V.B., then the probability vectors $\boldsymbol{f}_t(x)$ satisfy the recurrence relations (see Koutras and Alexandrou (1995), Fu (1996) and Koutras (1997))

$$\begin{aligned} &\boldsymbol{f}_t(0) = \boldsymbol{f}_{t-1}(0)A_t(0), \\ &\boldsymbol{f}_t(x) = \boldsymbol{f}_{t-1}(x)A_t(x) + \boldsymbol{f}_{t-1}(x-1)B_t(x-1), \qquad 1 \leq x \leq \ell_n,\ 1 \leq t \leq n. \end{aligned}$$

In addition, we introduce the following matrix

(iii) the return states matrix $C_t(x) = (\Pr(Y_t = U_{x-1,j} \mid Y_{t-1} = U_{x,i}))_{s\times s}$.

If X_n is an M.V.R., then the probability vectors $\boldsymbol{f}_t(x)$ satisfy the recurrence relations (see Han and Aki (1999))

$$\begin{aligned} &\boldsymbol{f}_t(x) = \boldsymbol{0}, \qquad x < 0 \quad \text{or} \quad x > \ell_t,\ 1 \leq t \leq n, \\ &\boldsymbol{f}_t(x) = \boldsymbol{f}_{t-1}(x)A_t(x) + \boldsymbol{f}_{t-1}(x-1)B_t(x-1) + \boldsymbol{f}_{t-1}(x+1)C_t(x+1), \\ &\hspace{22em} 0 \leq x \leq \ell_n,\ 1 \leq t \leq n. \end{aligned}$$

In both types of X_n, making use of the recursive schemes, we can easily evaluate the $\boldsymbol{f}_n(x)$. Therefore, the probability distribution function of X_n can be obtained. Remark that $A_t(x)+B_t(x)$ is a stochastic matrix if X_n is an M.V.B. and $A_t(x)+B_t(x)+C_t(x)$ is a stochastic matrix if X_n is an M.V.R.

As already stated in Introduction, we should make extensive use of the Markov chain imbedding method in order to study the joint distribution of $(X_n, F_{n,0}, S_{n,1}, \dots, S_{n,m})$ and the joint distribution of $(T_r, F_{T_r,0}, S_{T_r,1}, \dots, S_{T_r,m})$.

3. M.V.B. case: Joint distributions of the numbers of patterns, successes and failures in n trials

Assume that the number X_n of occurrences of $\mathcal{E}$ in n trials is an M.V.B. In this section, we consider the joint distribution of $(X_n, F_{n,0}, S_{n,1}, \dots, S_{n,m})$.

To begin with, we introduce the following transition probability matrices

$$\begin{aligned}
A_{t,0}(x,\boldsymbol{y}) &= (\Pr(Y_t = U_{x,i'}, \boldsymbol{S}_t = \boldsymbol{y} \mid Y_{t-1} = U_{x,i}, \boldsymbol{S}_{t-1} = \boldsymbol{y}))_{s\times s},\\
A_{t,j}(x,\boldsymbol{y}) &= (\Pr(Y_t = U_{x,i'}, \boldsymbol{S}_t = \boldsymbol{y}+\boldsymbol{e}_j \mid Y_{t-1} = U_{x,i}, \boldsymbol{S}_{t-1} = \boldsymbol{y}))_{s\times s},\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j=1,2,\dots,m,\\
B_{t,0}(x,\boldsymbol{y}) &= (\Pr(Y_t = U_{x+1,i'}, \boldsymbol{S}_t = \boldsymbol{y} \mid Y_{t-1} = U_{x,i}, \boldsymbol{S}_{t-1} = \boldsymbol{y}))_{s\times s},\\
B_{t,j}(x,\boldsymbol{y}) &= (\Pr(Y_t = U_{x+1,i'}, \boldsymbol{S}_t = \boldsymbol{y}+\boldsymbol{e}_j \mid Y_{t-1} = U_{x,i}, \boldsymbol{S}_{t-1} = \boldsymbol{y}))_{s\times s},\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j=1,2,\dots,m,
\end{aligned}$$

and the probability vectors

$$\begin{aligned}
\boldsymbol{f}_t(x,\boldsymbol{y}) = (\Pr(Y_t = U_{x,0}, \boldsymbol{S}_t = \boldsymbol{y}), \Pr(Y_t = U_{x,1}, \boldsymbol{S}_t = \boldsymbol{y}), \dots,&\\
\Pr(Y_t = U_{x,s-1}, \boldsymbol{S}_t = \boldsymbol{y})),& \qquad t\ge 1,
\end{aligned}$$

where, we denote the j-th unit vector of $\mathcal{R}^s$ by $\boldsymbol{e}_j$. Manifestly

$$f_n(x,\boldsymbol{y}) = \Pr(X_n = x, \boldsymbol{S}_n = y) = \sum_{i=0}^{s-1} \Pr(Y_n = U_{x,i}, \boldsymbol{S}_n = \boldsymbol{y}) = \boldsymbol{f}_n(x,\boldsymbol{y})\mathbf{1}'.$$

Therefore, we can obtain the joint probability distribution of $(X_n, S_{n,1}, \dots, S_{n,m})$ by evaluating the $\boldsymbol{f}_n(x,\boldsymbol{y})$. The next theorem provides a method for the evaluation of the joint probability distribution of $(X_n, S_{n,1}, \dots, S_{n,m})$.

THEOREM 3.1. *The probability vectors $\boldsymbol{f}_t(x,\boldsymbol{y})$, $(t\ge 2)$ satisfy the recurrence relations*

$$\begin{aligned}
\text{(3.1)}\qquad \boldsymbol{f}_t(x,\boldsymbol{y}) &= \boldsymbol{f}_{t-1}(x,\boldsymbol{y})A_{t,0}(x,\boldsymbol{y}) + \sum_{j=1}^{m} \boldsymbol{f}_{t-1}(x,\boldsymbol{y}-\boldsymbol{e}_j)A_{t,j}(x,\boldsymbol{y}-\boldsymbol{e}_j)\\
&\quad + \boldsymbol{f}_{t-1}(x-1,\boldsymbol{y})B_{t,0}(x-1,\boldsymbol{y})\\
&\quad + \sum_{j=1}^{m} \boldsymbol{f}_{t-1}(x-1,\boldsymbol{y}-\boldsymbol{e}_j)B_{t,j}(x-1,\boldsymbol{y}-\boldsymbol{e}_j)\\
&\qquad \textit{for}\quad t\ge 2,\ 0\le x\le \ell_t \quad\textit{and}\quad 0\le y_1, y_2, \dots, y_m \le t - xn_0,
\end{aligned}$$

with initial conditions

$$\begin{aligned}
\boldsymbol{f}_1(0,\boldsymbol{y}) = (\Pr(Y_1 = U_{0,0}, \boldsymbol{S}_1 = \boldsymbol{y}), \Pr(Y_1 = U_{0,1}, \boldsymbol{S}_1 = \boldsymbol{y}), \dots,&\\
\Pr(Y_1 = U_{0,s-1}, \boldsymbol{S}_1 = \boldsymbol{y})),&
\end{aligned}$$

for $y_1, y_2, \ldots, y_m = 0, 1$. *In addition, the joint probability distribution function of* $(X_n, S_{n,1}, \ldots, S_{n,m})$ *is given by*

$$\Pr(X_n = x, \boldsymbol{S}_n = \boldsymbol{y}) = \boldsymbol{f}_n(x, \boldsymbol{y})\mathbf{1}', \quad 0 \le x \le \ell_n, \ 0 \le y_1, y_2, \ldots, y_m \le n - xn_0. \tag{3.2}$$

PROOF. Note that the event $\{Y_t = U_{x,j}, \boldsymbol{S}_t = \boldsymbol{y}\}$ implies the occurrence of one of the events

$$\bigcup_{j=0}^{s-1} \left(\{Y_{t-1} = U_{x,j}, \boldsymbol{S}_{t-1} = \boldsymbol{y}\} \cup \bigcup_{i=1}^{m} \{Y_{t-1} = U_{x,j}, \boldsymbol{S}_{t-1} = \boldsymbol{y} - \boldsymbol{e}_i\} \right)$$

or

$$\bigcup_{j=0}^{s-1} \left(\{Y_{t-1} = U_{x-1,j}, \boldsymbol{S}_{t-1} = \boldsymbol{y}\} \cup \bigcup_{i=1}^{m} \{Y_{t-1} = U_{x-1,j}, \boldsymbol{S}_{t-1} = \boldsymbol{y} - \boldsymbol{e}_i\} \right).$$

The recurrences (3.1) are immediate consequences of the total probability theorem. It is easy to check the equation (3.2) from the following formula,

$$\Pr(X_n = x, \boldsymbol{S}_n = \boldsymbol{y}) = \Pr(Y_n \in \boldsymbol{U}_x, \boldsymbol{S}_n = \boldsymbol{y}) = \sum_{j=0}^{s-1} \Pr(Y_n = U_{x,j}, \boldsymbol{S}_n = \boldsymbol{y}). \qquad \square$$

Remark 3.1. In the special case where the matrices $A_{t,j}(x, \boldsymbol{y})$ and $B_{t,j}(x, \boldsymbol{y})$ do not depend on $\boldsymbol{y}$, that is,

$$A_{t,j}(x, \boldsymbol{y}) = A_{t,j}(x), \qquad B_{t,j}(x, \boldsymbol{y}) = B_{t,j}(x), \qquad \text{for all} \quad \boldsymbol{y},$$

then the matrices $\sum_{j=0}^{m} A_{t,j}(x)$ and $\sum_{j=0}^{m} B_{t,j}(x)$ are equal to the within state transition matrix $A_t(x)$ and the between state transition matrix $B_t(x)$, respectively. Therefore, the matrix $\sum_{j=0}^{m}(A_{t,j}(x) + B_{t,j}(x))$ is a stochastic matrix.

If $A_{t,j}(x, \boldsymbol{y}) = A_{t,j}$ and $B_{t,j}(x, \boldsymbol{y}) = B_{t,j}$ $(j = 0, 1, \ldots, m)$ for all $x, \boldsymbol{y}$, then the joint pgf $\phi_n(u, \boldsymbol{v})$ defined as (2.1) can be expressed as a product in the following way.

THEOREM 3.2. *If* $A_{t,j}(x, \boldsymbol{y}) = A_{t,j}$, $B_{t,j}(x, \boldsymbol{y}) = B_{t,j}$ $(j = 0, 1, \ldots, m)$ *for all* $x, \boldsymbol{y}$, *then the joint pgf* $\phi_n(u, \boldsymbol{v})$ *of* $(X_n, S_{n,1}, \ldots, S_{n,m})$ *can be expressed as*

$$\phi_n(u, \boldsymbol{v}) = \boldsymbol{a}(\boldsymbol{v}) \prod_{t=2}^{n} \left[A_{t,0} + uB_{t,0} + \sum_{j=1}^{m} v_j (A_{t,j} + uB_{t,j}) \right] \mathbf{1}' \tag{3.3}$$

where,

$$\boldsymbol{a}(\boldsymbol{v}) = \sum_{\boldsymbol{y}=\boldsymbol{0}}^{\boldsymbol{1}} \boldsymbol{f}_1(0, \boldsymbol{y}) \boldsymbol{v}^{\boldsymbol{y}}. \tag{3.4}$$

PROOF. Introducing the vector generating functions

$$\boldsymbol{\phi}_t(u, \boldsymbol{v}) = \sum_{x=0}^{\ell_t} \sum_{\boldsymbol{y}=0}^{t-xn_0} \boldsymbol{f}_t(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}},$$

we can write

$$\phi_t(u, \boldsymbol{v}) = \boldsymbol{\phi}_t(u, \boldsymbol{v})\mathbf{1}'.$$

By summing both sides of the equation (3.1) after multiplying by $u^x \boldsymbol{v}^{\boldsymbol{y}}$, $\boldsymbol{\phi}_t(u, \boldsymbol{v})$ can be expressed as

$$\begin{aligned}\boldsymbol{\phi}_t(u, \boldsymbol{v}) &= \sum_{x=0}^{\ell_t} \sum_{\boldsymbol{y}=\boldsymbol{0}}^{t-1-xn_0} \boldsymbol{f}_{t-1}(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} A_{t,0} + \sum_{j=1}^{m} v_j \sum_{x=0}^{\ell_t} \sum_{\boldsymbol{y}=\boldsymbol{0}}^{t-1-xn_0} \boldsymbol{f}_{t-1}(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} A_{t,j} \\ &\quad + u \sum_{x=0}^{\ell_t - 1} \sum_{\boldsymbol{y}=\boldsymbol{0}}^{t-1-xn_0} \boldsymbol{f}_{t-1}(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} B_{t,0} \\ &\quad + u \sum_{j=1}^{m} v_j \sum_{x=0}^{\ell_t - 1} \sum_{\boldsymbol{y}=\boldsymbol{0}}^{t-1-xn_0} \boldsymbol{f}_{t-1}(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} B_{t,j}.\end{aligned}$$

We consider the two possible cases $\ell_t = \ell_{t-1}$ and $\ell_t = \ell_{t-1} + 1$ separately. If $\ell_t = \ell_{t-1}$, we should note the identity

$$\boldsymbol{f}_{t-1}(\ell_{t-1}, \boldsymbol{y}) B_{t,j} = \boldsymbol{0}, \qquad \text{for} \quad j = 0, 1, \dots, m.$$

If $\ell_t = \ell_{t-1} + 1$, we should note the identity

$$\boldsymbol{f}_{t-1}(\ell_t, \boldsymbol{y}) A_{t,j} = \boldsymbol{0}, \qquad \text{for} \quad j = 0, 1, \dots, m.$$

Hence, in both cases we have

$$\boldsymbol{\phi}_t(u, \boldsymbol{v}) = \boldsymbol{\phi}_{t-1}(u, \boldsymbol{v}) \left[A_{t,0} + uB_{t,0} + \sum_{j=1}^{m} v_j (A_{t,j} + uB_{t,j}) \right].$$

If we take into account that

$$\boldsymbol{\phi}_1(u, \boldsymbol{v}) = \sum_{x=0}^{\ell_1} \sum_{\boldsymbol{y}=\boldsymbol{0}}^{\boldsymbol{1}} \boldsymbol{f}_1(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} = \sum_{\boldsymbol{y}=\boldsymbol{0}}^{\boldsymbol{1}} \boldsymbol{f}_1(0, \boldsymbol{y}) \boldsymbol{v}^{\boldsymbol{y}} = \boldsymbol{a}(\boldsymbol{v}),$$

the proof is completed. □

For the homogeneous case (i.e. $A_{t,j} = A_j$, $B_{t,j} = B_j$, $j = 0, 1, \dots, m$), the double generating function $\Phi(u, \boldsymbol{v}; w)$ defined as (2.2) takes more compact form.

THEOREM 3.3. *If $A_{t,j}(x, \boldsymbol{y}) = A_j$, $B_{t,j}(x, \boldsymbol{y}) = B_j$ $(j = 0, 1, \dots, m)$ for all $x, \boldsymbol{y}$ and $t \geq 1$, then the double generating function of $(X_n, S_{n,1}, \dots, S_{n,m})$ is given by*

$$\Phi(u, \boldsymbol{v}; w) = w\boldsymbol{a}(\boldsymbol{v}) \left[I - w \left(A_0 + uB_0 + \sum_{j=1}^{m} v_j (A_j + uB_j) \right) \right]^{-1} \mathbf{1}', \tag{3.5}$$

where $\boldsymbol{a}(\boldsymbol{v})$ is given by (3.4).

PROOF. From the equation (3.3) under the conditions that $A_{t,j}(x,\boldsymbol{y}) = A_j$, $B_{t,j}(x,\boldsymbol{y}) = B_j$, $(j = 0, 1, \ldots, m)$, the joint pgf $\phi_n(u, \boldsymbol{v})$ can be expressed as

$$\phi_n(u,\boldsymbol{v}) = \boldsymbol{a}(\boldsymbol{v}) \left[A_0 + uB_0 + \sum_{j=1}^{m} v_j(A_j + uB_j) \right]^{n-1} \mathbf{1}'. \tag{3.6}$$

Making use of the formula

$$\sum_{n=0}^{\infty} \left[A_0 + uB_0 + \sum_{j=1}^{m} v_j(A_j + uB_j) \right]^{n} w^n$$
$$= \left[I - w\left(A_0 + uB_0 + \sum_{j=1}^{m} v_j(A_j + uB_j) \right) \right]^{-1},$$

we have the desired result. □

Let $F_{n,0}$ be the number of the outcome "0". Then we can obtain the joint pgf $\psi_n(u, v_0, \boldsymbol{v})$ of $(X_n, F_{n,0}, S_{n,1}, \ldots, S_{n,m})$ through $\phi_n(u, \boldsymbol{v})$:

$$\begin{aligned}
\psi_n(u, v_0, \boldsymbol{v}) &= E(u^{X_n} v_0^{F_{n,0}} v_1^{S_{n,1}} \cdots v_m^{S_{n,m}}) \\
&= E(u^{X_n} v_0^{n-S_{n,1}-S_{n,2}-\cdots-S_{n,m}} v_1^{S_{n,1}} \cdots v_m^{S_{n,m}}) \\
&= v_0^n \phi_n(u, v_1/v_0, \ldots, v_m/v_0).
\end{aligned}$$

Therefore the double generating function

$$\Psi(u, v_0, \boldsymbol{v}; w) = \sum_{n=1}^{\infty} \psi_n(u, v_0, \boldsymbol{v}) w^n$$

takes the form $\Psi(u, v_0, \boldsymbol{v}; w) = \Phi(u, v_1/v_0, \ldots, v_m/v_0; wv_0)$. More specifically, in terms of (3.5), we have the following result.

THEOREM 3.4. *If $A_{t,j}(x,\boldsymbol{y}) = A_j$, $B_{t,j}(x,\boldsymbol{y}) = B_j$ $(j = 0, 1, \ldots, m)$ for all $x, \boldsymbol{y}$ and $t \geq 1$, then the double generating function $\Psi(u, v_0, \boldsymbol{v}; w)$ of $(X_n, F_{n,0}, S_{n,1}, \ldots, S_{n,m})$ can be expressed as*

$$\Psi(u, v_0, \boldsymbol{v}; w) = wv_0 \boldsymbol{a}(v_1/v_0, \ldots, v_m/v_0) \left[I - w \sum_{j=0}^{m} v_j(A_j + uB_j) \right]^{-1} \mathbf{1}', \tag{3.7}$$

where $\boldsymbol{a}(\boldsymbol{v})$ is given by (3.4).

As by-products, we can derive some interesting formulae from the equation (3.7). The double generating functions of $(X_n, F_{n,0})$, $(X_n, S_{n,i})$ $(i = 1, 2, \ldots, m)$ are given by

$$\sum_{n=1}^{\infty} E(u^{X_n} v_0^{F_{n,0}}) w^n = \Psi(u, v_0, \boldsymbol{v}; w) \bigg|_{v_1 = v_2 = \cdots = v_m = 1}$$

$$= wv_0\boldsymbol{a}(1/v_0,\ldots,1/v_0)$$
$$\times\left[I - w\left(v_0(A_0+uB_0)+\sum_{j=1}^{m}(A_j+uB_j)\right)\right]^{-1}\mathbf{1}',$$
$$\sum_{n=1}^{\infty}E(u^{X_n}v_i^{S_{n,i}})w^n = \Psi(u,v_0,\boldsymbol{v};w)\Big|_{v_0=v_1=\cdots=v_{i-1}=1,v_{i+1}=\cdots=v_m=1}$$
$$= w\boldsymbol{a}(\underbrace{1,\ldots,1}_{i-1},v_i,\underbrace{1,\ldots,1}_{m-i})$$
$$\times\left[I - w\left(v_i(A_i+uB_i)+\sum_{j\neq i}(A_j+uB_j)\right)\right]^{-1}\mathbf{1}',$$
$$i=1,2,\ldots,m.$$

The double generating function of X_n is

$$\sum_{n=1}^{\infty}E(w^{X_n})u^n = \Psi(u,v_0,\boldsymbol{v};w)\mid_{v_0=v_1=\cdots=v_m=1} = w\boldsymbol{a}(\mathbf{1})[I-w(A+uB)]^{-1}\mathbf{1}',$$

where, $A=\sum_{j=0}^{m}A_j$, $B=\sum_{j=0}^{m}B_j$ and $\boldsymbol{a}(\mathbf{1})=\boldsymbol{a}(\boldsymbol{v})\mid_{v_1=\cdots=v_m=1}$ (for a direct proof of an analogous result see Koutras and Alexandrou (1995)).

THEOREM 3.5. *If $A_{t,j}(x,\boldsymbol{y})=A_j$, $B_{t,j}(x,\boldsymbol{y})=B_j$ $(j=0,1,\ldots,m)$ for all $x,\boldsymbol{y}$ and $t\geq 1$, then the expected value of X_n and its generating function are given by*

$$M_{X_n}(w)=\sum_{n=1}^{\infty}E(X_n)w^n=\frac{w^2}{1-w}\boldsymbol{a}(\mathbf{1})[I-w(A+B)]^{-1}B\mathbf{1}',$$
$$E(X_n)=\boldsymbol{a}(\mathbf{1})\sum_{i=1}^{n-1}(A+B)^{i-1}B\mathbf{1}'.$$

where, $\boldsymbol{a}(\mathbf{1})=\boldsymbol{a}(\boldsymbol{v})\mid_{v_1=v_2=\cdots=v_m=1}$, $A=\sum_{j=0}^{m}A_j$ and $B=\sum_{j=0}^{m}B_j$.

PROOF. Note that

$$\sum_{n=1}^{\infty}E(X_n)w^n=\frac{\partial}{\partial u}\Psi(u,v_0,\boldsymbol{v};w)\Big|_{u=v_0=v_1=\cdots=v_m=1}$$

and making use of

$$\frac{d}{dz}(U-zV)^{-1}=(U-zV)^{-1}V(U-zV)^{-1},$$

we have the first conclusion of the theorem. By expanding $M_{X_n}(w)$ in a power series of w, we have the expected value of X_n immediately. □

The expressions in Theorem 3.5 are easily shown to be consistent with similar expressions given in Koutras and Alexandrou (1995), Han and Aki (1999) and Chadjiconstantinidis *et al.* (2000). The formulae for the variance of X_n, the covariance between X_n and $F_{n,0}$ and the covariances between X_n and $S_{n,i}$ $(i=1,2,\ldots,m)$ can be derived from the derivatives of the double generating function $\Psi(u,v_0,\boldsymbol{v};w)$, however, their expressions are not very attractive.

4. M.V.B. case: Joint distributions of the waiting time and the numbers of successes and failures

Assume that the number X_n of occurrences of $\mathcal{E}$ in n trials is an M.V.B. In this section, we will consider the joint distribution of $(T_r, F_{T_r,0}, S_{T_r,1}, \ldots, S_{T_r,m})$.

THEOREM 4.1. *The joint probability distribution function of* $(T_r, S_{T_r,1}, \ldots, S_{T_r,m})$ *can be expressed as*

$$(4.1) \qquad h_r(n, \boldsymbol{y}) = \sum_{i=1}^{s} \left[\beta_{i,0}(n; r, y) \boldsymbol{f}_{n-1}(r-1, \boldsymbol{y}) + \sum_{j=1}^{m} \beta_{i,j}(n; r, y) \boldsymbol{f}_{n-1}(r-1, \boldsymbol{y} - \boldsymbol{e}_j) \right] \boldsymbol{e}_i',$$

$$n \geq 2, \quad 0 \leq y_1, y_2, \ldots, y_m \leq n,$$

$$h_r(1, \boldsymbol{y}) = 0, \qquad y_1, y_2, \ldots, y_m = 0, 1,$$

where,

$$\beta_{i,0}(n; r, \boldsymbol{y}) = \boldsymbol{e}_i B_{n,0}(r-1, \boldsymbol{y} - \boldsymbol{e}_j)\boldsymbol{1}',$$
$$\beta_{i,j}(n; r, \boldsymbol{y}) = \boldsymbol{e}_i B_{n,j}(r-1, \boldsymbol{y} - \boldsymbol{e}_j)\boldsymbol{1}', \qquad j = 1, 2, \ldots, m.$$

PROOF. Note that

$$h_r(n, \boldsymbol{y}) = \Pr(T_r = n, \boldsymbol{S}_{T_r} = \boldsymbol{y}) = \Pr(T_r = n, \boldsymbol{S}_n = \boldsymbol{y}),$$

which is equivalent to

$$h_r(n, \boldsymbol{y}) = \Pr(X_n = r, X_{n-1} = r-1, \boldsymbol{S}_n = \boldsymbol{y}) = \Pr(Y_n \in \boldsymbol{U}_r, Y_{n-1} \in \boldsymbol{U}_{r-1}, \boldsymbol{S}_n = \boldsymbol{y}).$$

Making use of the further decomposition of the event $\{Y_n \in \boldsymbol{U}_r, Y_{n-1} \in \boldsymbol{U}_{r-1}, \boldsymbol{S}_n = \boldsymbol{y}\}$, we have

$$\begin{aligned}
h_r(n, \boldsymbol{y}) &= \Pr(Y_n \in \boldsymbol{U}_r, Y_{n-1} \in \boldsymbol{U}_{r-1}, \boldsymbol{S}_n = \boldsymbol{y}) \\
&= \sum_{i=0}^{s-1} \Pr(Y_n \in \boldsymbol{U}_r, \boldsymbol{S}_n = \boldsymbol{y} \mid Y_{n-1} = U_{r-1,i}, \boldsymbol{S}_{n-1} = \boldsymbol{y}) \\
&\quad \times \Pr(Y_{n-1} = U_{r-1,i}, \boldsymbol{S}_{n-1} = \boldsymbol{y}) \\
&\quad + \sum_{i=0}^{s-1} \sum_{j=1}^{m} \Pr(Y_n \in \boldsymbol{U}_r, \boldsymbol{S}_n = \boldsymbol{y} \mid Y_{n-1} = U_{r-1,i}, \boldsymbol{S}_{n-1} = \boldsymbol{y} - \boldsymbol{e}_j) \\
&\quad \times \Pr(Y_{n-1} = U_{r-1,i}, \boldsymbol{S}_{n-1} = \boldsymbol{y} - \boldsymbol{e}_j) \\
&= \sum_{i=0}^{s-1} \boldsymbol{e}_{i+1} B_{n,0}(r-1, \boldsymbol{y})\boldsymbol{1}' \Pr(Y_{n-1} = U_{r-1,i}, \boldsymbol{S}_{n-1} = \boldsymbol{y}) \\
&\quad + \sum_{i=0}^{s-1} \sum_{j=1}^{m} \boldsymbol{e}_{i+1} B_{n,j}(r-1, \boldsymbol{y} - \boldsymbol{e}_j)\boldsymbol{1}' \Pr(Y_{n-1} = U_{r-1,i}, \boldsymbol{S}_{n-1} = \boldsymbol{y} - \boldsymbol{e}_j) \\
&= \sum_{i=1}^{s} \left[\beta_{i,0}(n; r, y) \boldsymbol{f}_{n-1}(r-1, \boldsymbol{y}) + \sum_{j=1}^{m} \beta_{i,j}(n; r, y) \boldsymbol{f}_{n-1}(r-1, \boldsymbol{y} - \boldsymbol{e}_j) \right] \boldsymbol{e}_i'.
\end{aligned}$$

Under the assumption that the length of pattern $\mathcal{E}$ is greater than 1, we have immediately $h_r(1, \boldsymbol{y}) = \Pr(T_r = 1, \boldsymbol{S}_1 = \boldsymbol{y}) = 0$. The proof is completed. □

In the case where $A_{t,j}(x, \boldsymbol{y}) = A_j$, $B_{t,j}(x, \boldsymbol{y}) = B_j$, $(j = 0, 1, \ldots, m)$, the double generating function $H(w, \boldsymbol{v}; u)$ defined as (2.4) takes more compact form.

THEOREM 4.2. *If $A_{t,j}(x, \boldsymbol{y}) = A_j$, $B_{t,j}(x, \boldsymbol{y}) = B_j$ $(j = 0, 1, \ldots, m)$ for all $x, \boldsymbol{y}$ and $t \geq 1$, then the double generating function of $(T_r, S_{T_r,1}, \ldots, S_{T_r,m})$ is given by*

$$(4.2) \qquad H(w, \boldsymbol{v}; u) = uw^2 \boldsymbol{a}(\boldsymbol{v}) \sum_{i=1}^{s} \left(\beta_{i,0} + \sum_{j=1}^{m} v_j \beta_{i,j} \right) \times \left[I - w \left(A_0 + uB_0 + \sum_{j'=1}^{m} v_{j'} (A_{j'} + uB_{j'}) \right) \right]^{-1} \boldsymbol{e}_i'$$

where, $\beta_{i,j} = \boldsymbol{e}_i B_j \boldsymbol{1}'$, $i = 1, \ldots, s$, $j = 0, 1, \ldots, m$, and $\boldsymbol{a}(\boldsymbol{v})$ is given by (3.4).

PROOF. From the equation (4.1) under the conditions that $A_{t,j}(x, \boldsymbol{y}) = A_j$, $B_{t,j}(x, \boldsymbol{y}) = B_j$, $(j = 0, 1, \ldots, m)$, the joint probability distribution function of $(T_r, S_{T_r,1}, \ldots, S_{T_r,m})$ is

$$(4.3) \qquad h_r(n, \boldsymbol{y}) = \sum_{i=1}^{s} \left[\beta_{i,0} \boldsymbol{f}_{n-1}(r-1, \boldsymbol{y}) + \sum_{j=1}^{m} \beta_{i,j} \boldsymbol{f}_{n-1}(r-1, \boldsymbol{y} - \boldsymbol{e}_j) \right] \boldsymbol{e}_i'.$$

A straightforward manipulation over (4.3) reveals that

$$\begin{aligned} H(w, \boldsymbol{v}; u) &= uw \sum_{i=1}^{s} \left(\beta_{i,0} + \sum_{j=1}^{m} v_j \beta_{i,j} \right) \sum_{n=1}^{\infty} \sum_{r=0}^{\infty} \sum_{\boldsymbol{y}=\boldsymbol{0}}^{\infty} \boldsymbol{f}_n(r, \boldsymbol{y}) u^r \boldsymbol{v}^{\boldsymbol{y}} w^n \boldsymbol{e}_i' \\ &= uw \sum_{i=1}^{s} \left(\beta_{i,0} + \sum_{j=1}^{m} v_j \beta_{i,j} \right) \sum_{n=1}^{\infty} \boldsymbol{\phi}_n(u, \boldsymbol{v}) w^n \boldsymbol{e}_i'. \end{aligned}$$

The proof is completed if we take into account that

$$\sum_{n=1}^{\infty} \boldsymbol{\phi}_n(u, \boldsymbol{v}) w^n = w \boldsymbol{a}(\boldsymbol{v}) \left[I - w \left(A_0 + uB_0 + \sum_{j'=1}^{m} v_{j'} (A_{j'} + uB_{j'}) \right) \right]^{-1}. \qquad \square$$

As described in the next theorem, expanding (4.2) in a power series of u, we can get the joint pgf $H_r(w, \boldsymbol{v})$ defined as (2.3).

THEOREM 4.3. *If $A_{t,j}(x, \boldsymbol{y}) = A_j$, $B_{t,j}(x, \boldsymbol{y}) = B_j$ $(j = 0, 1, \ldots, m)$ for all $x, \boldsymbol{y}$ and $t \geq 1$, then the joint pgf of $(T_r, S_{T_r,1}, \ldots, S_{T_r,m})$ is given by*

$$(4.4) \qquad H_r(w, \boldsymbol{v}) = w^{r+1} \boldsymbol{a}(\boldsymbol{v}) \sum_{i=1}^{s} \left(\beta_{i,0} + \sum_{j=1}^{m} v_j \beta_{i,j} \right)$$

$$\times\left[\left(I-w\left(A_0+\sum_{j'=1}^{m}v_{j'}A_{j'}\right)\right)^{-1}\left(B_0+\sum_{j'=1}^{m}v_{j'}B_{j'}\right)\right]^{r-1}$$
$$\times\left(I-w\left(A_0+\sum_{j'=1}^{m}v_{j'}A_{j'}\right)\right)^{-1}\boldsymbol{e}_i',\qquad r\geq 1,$$

where $\boldsymbol{a}(\boldsymbol{v})$ *is given by* (3.4).

PROOF. Since

$$I-w\left(\left(A_0+\sum_{j'=1}^{m}v_{j'}A_{j'}\right)+u\left(B_0+\sum_{j'=1}^{m}v_{j'}B_{j'}\right)\right)$$
$$=\left(I-w\left(A_0+\sum_{j'=1}^{m}v_{j'}A_{j'}\right)\right)$$
$$\times\left[I-wu\left(I-w\left(A_0+\sum_{j'=1}^{m}v_{j'}A_{j'}\right)\right)^{-1}\left(B_0+\sum_{j'=1}^{m}v_{j'}B_{j'}\right)\right],$$

it follows that

$$\left[I-w\left(\left(A_0+\sum_{j'=1}^{m}v_{j'}A_{j'}\right)+u\left(B_0+\sum_{j'=1}^{m}v_{j'}B_{j'}\right)\right)\right]^{-1}$$
$$=\sum_{j=0}^{\infty}\left[\left(I-w\left(A_0+\sum_{j'=1}^{m}v_{j'}A_{j'}\right)\right)^{-1}\left(B_0+\sum_{j'=1}^{m}v_{j'}B_{j'}\right)\right]^{j}$$
$$\times\left(I-w\left(A_0+\sum_{j'=1}^{m}v_{j'}A_{j'}\right)\right)^{-1}$$
$$\times(wu)^j.$$

Hence, we can write $H(w,\boldsymbol{v};u)$ as

$$H(w,\boldsymbol{v};u)=uw^2\boldsymbol{a}(\boldsymbol{v})\sum_{i=1}^{s}\left(\beta_{i,0}+\sum_{j=1}^{m}v_j\beta_{i,j}\right)$$
$$\times\sum_{j=0}^{\infty}\left[\left(I-w\left(A_0+\sum_{j'=1}^{m}v_{j'}A_{j'}\right)\right)^{-1}\left(B_0+\sum_{j'=1}^{m}v_{j'}B_{j'}\right)\right]^{j}$$
$$\times\left(I-w\left(A_0+\sum_{j'=1}^{m}v_{j'}A_{j'}\right)\right)^{-1}$$
$$\times(wu)^j\boldsymbol{e}_i'$$

or equivalently in the more interesting form

$$\begin{aligned} H(w,\boldsymbol{v};u) &= \sum_{r=1}^{\infty} w^{r+1}\boldsymbol{a}(\boldsymbol{v}) \sum_{i=1}^{s}\left(\beta_{i,0}+\sum_{j=1}^{m} v_j\beta_{i,j}\right) \\ &\quad\times \left[\left(I - w\left(A_0 + \sum_{j'=1}^{m} v_{j'}A_{j'}\right)\right)^{-1}\left(B_0 + \sum_{j'=1}^{m} v_{j'}B_{j'}\right)\right]^{r-1} \\ &\quad\times \left(I - w\left(A_0 + \sum_{j'=1}^{m} v_{j'}A_{j'}\right)\right)^{-1} \\ &\quad\times \boldsymbol{e}_i' u^r, \end{aligned}$$

which manifestly yields the desired result. □

In the special case $r = 1$, from the formula (4.4), we can derive the joint pgf $H_1(w,\boldsymbol{v})$ related to the first occurrence of $\mathcal{E}$. More quickly, by exploiting the formula

$$H_1(w,\boldsymbol{v}) = \left[\frac{1}{u}H(w,\boldsymbol{v};u)\right]_{u=0},$$

we can also derive the joint pgf $H_1(w,\boldsymbol{v})$ of $(T_1, S_{T_1,1},\dots,S_{T_1,m})$.

COROLLARY 4.1. *If $A_{t,j}(x,\boldsymbol{y}) = A_j$, $B_{t,j}(x,\boldsymbol{y}) = B_j$ $(j = 0,1,\dots,m)$ for all $x,\boldsymbol{y}$ and $t \geq 1$, then the joint pgf of $(T_1, S_{T_1,1},\dots,S_{T_1,m})$ is given by*

$$H_1(w,\boldsymbol{v}) = w^2\boldsymbol{a}(\boldsymbol{v})\sum_{i=1}^{s}\left(\beta_{i,0}+\sum_{j=1}^{m} v_j\beta_{i,j}\right)\left[I - w\left(A_0 + \sum_{j'=1}^{m} v_{j'}A_{j'}\right)\right]^{-1}\boldsymbol{e}_i',$$

where, $\beta_{i,j} = \boldsymbol{e}_iB_j\boldsymbol{1}'$, $i = 1,\dots,s$, $j = 0,1,\dots,m$, and $\boldsymbol{a}(\boldsymbol{v})$ is given by (3.4).

In a similar fashion as in the conclusion of Section 3, we can easily establish formulae for the joint pgf of $(T_r, F_{T_r,0}, S_{T_r,1},\dots,S_{T_r,m})$. Through $H_r(w,\boldsymbol{v})$, the joint pgf $G_r(w,v_0,\boldsymbol{v})$ of $(T_r, F_{T_r,0}, S_{T_r,1},\dots,S_{T_r,m})$ can be expressed as

$$\begin{aligned} G_r(w,v_0,\boldsymbol{v}) &= E(w^{T_r}v_0^{F_{T_r,0}}v_1^{S_{T_r,1}}\cdots v_m^{S_{T_r,m}}) \\ &= E(w^{T_r}v_0^{T_r - S_{T_r,1}-S_{T_r,2}-\cdots-S_{T_r,m}}v_1^{S_{T_r,1}}\cdots v_m^{S_{T_r,m}}) \\ &= H_r(wv_0, v_1/v_0, v_2/v_0,\dots,v_m/v_0). \end{aligned}$$

Through $H(w,\boldsymbol{v};u)$, the double generating function of $(T_r, F_{T_r,0}, S_{T_r,1},\dots,S_{T_r,m})$ is also expressed as

$$\begin{aligned} G(w,v_0,\boldsymbol{v};u) &= \sum_{r=1}^{\infty} G_r(w,v_0,\boldsymbol{v})u^r \\ &= \sum_{r=1}^{\infty} H_r(wv_0, v_1/v_0, v_2/v_0,\dots,v_m/v_0)u^r \\ &= H(wv_0, v_1/v_0,\dots,v_m/v_0;u). \end{aligned}$$

More specifically, in terms of (4.2) and (4.4), we have the following result.

THEOREM 4.4. *Assume that* $A_{t,j}(x,\boldsymbol{y}) = A_j$, $B_{t,j}(x,\boldsymbol{y}) = B_j$ $(j = 0,1,\ldots,m)$ *for all* $x, \boldsymbol{y}$ *and* $t \geq 1$. *Then the double generating function* $G(w, v_0, \boldsymbol{v}; u)$ *of* $(T_r, F_{T_r,0}, S_{T_r,1}, \ldots, S_{T_r,m})$ *is given by*

$$G(w,v_0,\boldsymbol{v};u) = uw^2 v_0 \boldsymbol{a}(v_1/v_0,\ldots,v_m/v_0)\sum_{i=1}^{s}\sum_{j=0}^{m} v_j\beta_{i,j} \times \left[I - w\sum_{j'=0}^{m} v_{j'}(A_{j'}+uB_{j'})\right]^{-1} \boldsymbol{e}_i', \tag{4.5}$$

and the joint pgf $G_r(w, v_0, \boldsymbol{v})$ *of* $(T_r, F_{T_r,0}, S_{T_r,1}, \ldots, S_{T_r,m})$ *is given by*

$$\begin{aligned} G_r(w,v_0,\boldsymbol{v}) &= w^{r+1} v_0 \boldsymbol{a}(v_1/v_0,\ldots,v_m/v_0) \\ &\quad \times \sum_{i=1}^{s}\sum_{j=0}^{m} v_j\beta_{i,j}\left[\left(I - w\sum_{j'=0}^{m} v_{j'}A_{j'}\right)^{-1}\left(\sum_{j'=0}^{m} v_{j'}B_{j'}\right)\right]^{r-1} \\ &\quad \times \left(I - w\sum_{j'=0}^{m} v_{j'}A_{j'}\right)^{-1}\boldsymbol{e}_i', \qquad r\geq 1, \end{aligned}$$

where, $\beta_{i,j} = \boldsymbol{e}_i B_j \boldsymbol{1}'$, $i = 1,\ldots,s$, $j = 0,1,\ldots,m$, *and* $\boldsymbol{a}(\boldsymbol{v})$ *is given by* (3.4).

As by-products, we can derive some interesting formulae from the equation (4.5). The double generating functions of $(T_r, F_{T_r,0})$, $(T_r, S_{T_r,i})$, $i = 1,2,\ldots,m$ are given by

$$\begin{aligned} \sum_{r=1}^{\infty} E(w^{T_r} v_0^{F_{T_r,0}})u^r &= G(w,v_0,\boldsymbol{v};u)\Big|_{v_1=\cdots=v_m=1} \\ &= uw^2 v_0 \boldsymbol{a}(1/v_0,\ldots,1/v_0)\sum_{j=1}^{s}\left(v_0\beta_{j,0} + \sum_{j'=1}^{m}\beta_{j,j'}\right) \\ &\quad \times \left[I - w\left(v_0(A+uB) + \sum_{j'=1}^{m}(A_{j'}+uB_{j'})\right)\right]^{-1}\boldsymbol{e}_j', \end{aligned}$$

$$\begin{aligned} \sum_{r=1}^{\infty} E(w^{T_r} v_i^{S_{T_r,i}})u^r &= G(w,v_0,\boldsymbol{v};u)\Big|_{v_0=v_1=\cdots v_{i-1}=1, v_{i+1}=\cdots v_m=1} \\ &= uw^2 \boldsymbol{a}(\underbrace{1,\ldots,1}_{i-1}, v_i, \underbrace{1,\ldots,1}_{m-i})\sum_{j=1}^{s}\left(v_i\beta_{j,i} + \sum_{j'\neq i}^{m}\beta_{j,j'}\right) \\ &\quad \times \left[I - w\left(v_i(A_i+uB_i) + \sum_{j'\neq i}^{m}(A_{j'}+uB_{j'})\right)\right]^{-1}\boldsymbol{e}_j', \\ &\qquad i = 1,2,\ldots,m. \end{aligned}$$

The double generating function of T_r can be expressed as

$$\sum_{r=1}^{\infty} E(w^{T_r})u^r = G(w, v_0, \boldsymbol{v}; u)\Big|_{v_0=v_1=\cdots=v_m=1} = uw^2\boldsymbol{a}(\mathbf{1})\sum_{j=1}^{s}\sum_{j'=0}^{m}\beta_{j,j'}\left[I - w(A+uB)\right]^{-1}\boldsymbol{e}'_j, \tag{4.6}$$

where, $A = \sum_{j=0}^{m} A_j$, $B = \sum_{j=0}^{m} B_j$ and $\boldsymbol{a}(\mathbf{1}) = \boldsymbol{a}(\boldsymbol{v})\mid_{v_1=\cdots=v_m=1}$ (for a direct proof of an analogous result see Koutras (1997)).

We can also establish the double generating functions of $F_{T_r,0}$, $S_{T_r,i}$ $(i = 1, 2, \ldots, m)$,

$$\begin{aligned}\sum_{r=1}^{\infty} E(v_0^{F_{T_r,0}})u^r &= G(w, v_0, \boldsymbol{v}; u)\Big|_{w=v_1=\cdots=v_m=1}\\ &= uv_0\boldsymbol{a}(1/v_0, \ldots, 1/v_0)\sum_{j=1}^{s}\left(v_0\beta_{j,0} + \sum_{j'=1}^{m}\beta_{j,j'}\right)\\ &\quad\times\left[I - \left(v_0(A+uB) + \sum_{j'=1}^{m}(A_{j'} + uB_{j'})\right)\right]^{-1}\boldsymbol{e}'_j,\end{aligned}$$

$$\begin{aligned}\sum_{r=1}^{\infty} E(v_i^{S_{T_r,i}})u^r &= G(w, v_0, \boldsymbol{v}; u)\Big|_{w=v_0=v_1=\cdots v_{i-1}=1, v_{i+1}=\cdots v_m=1}\\ &= u\boldsymbol{a}(\underbrace{1, \ldots, 1}_{i-1}, v_i, \underbrace{1, \ldots, 1}_{m-i})\sum_{j=1}^{s}\left(v_i\beta_{j,i} + \sum_{j'\neq i}^{m}\beta_{j,j'}\right)\\ &\quad\times\left[I - \left(v_i(A_i+uB_i) + \sum_{j'\neq i}^{m}(A_{j'} + uB_{j'})\right)\right]^{-1}\boldsymbol{e}'_j,\\ &\qquad\qquad i = 1, 2, \ldots, m.\end{aligned}$$

By differentiating the above expressions, we can establish the formulae for the expected values of T_r, $F_{T_r,0}$, $S_{T_r,i}$ $(i = 1, 2, \ldots, m)$, their variances, the covariance between T_r and $F_{T_r,0}$ and the covariances between T_r and $S_{T_r,i}$ $(i = 1, 2, \ldots, m)$. For example, from the formula (4.6), we get

$$\begin{aligned}\sum_{r=1}^{\infty} E(T_r)u^r &= 2u\boldsymbol{a}(\mathbf{1})\sum_{j=1}^{s}\sum_{j'=0}^{m}\beta_{j,j'}(I - A - uB)^{-1}\boldsymbol{e}'_j\\ &\quad + u\boldsymbol{a}(\mathbf{1})\sum_{j=1}^{s}\sum_{j'=0}^{m}\beta_{j,j'}(I - A - uB)^{-1}(A+uB)(I - A - uB)^{-1}\boldsymbol{e}'_j.\end{aligned}$$

Aki and Hirano (2000) considered the distributions of the numbers of non-overlapping occurrences of "1" runs of length k until the n-th occurrence of "1" in a sequence of $\{0, 1\}$-valued random variables. They called the distribution "the generalized binomial distribution of order $(k-1)$". Through slight modifications of our methods and results in this secton, we can deal with the problems as well.

Closing, we mention that Inoue and Aki (2002) studied the generalized waiting time problem for the first occurrence of a pattern in a sequence obtained by Pólya's urn scheme. They present a completely different method, which is based on the methods of conditional probability generating functions and a notion of truncation.

5. M.V.R. case: Joint distributions of the numbers of patterns, successes and failures in n trials

We assume that the number X_n of occurrences of $\mathcal{E}$ in n trials is an M.V.R. This section will deal with the joint distribution of $(X_n, F_{n,0}, S_{n,1}, \dots, S_{n,m})$. In this section, we will use the same notations and terminology as in Section 3. In addition to these, we introduce the $s \times s$ transition probability matrices for the Markov chain $\{Y_t, t \geq 0\}$.

$$\begin{aligned} C_{t,0}(x, \boldsymbol{y}) &= (\Pr(Y_t = U_{x-1,i'}, \boldsymbol{S}_t = \boldsymbol{y} \mid Y_{t-1} = U_{x,i}, \boldsymbol{S}_{t-1} = \boldsymbol{y}))_{s\times s}, \\ C_{t,j}(x, \boldsymbol{y}) &= (\Pr(Y_t = U_{x-1,i'}, \boldsymbol{S}_t = \boldsymbol{y} + \boldsymbol{e}_j \mid Y_{t-1} = U_{x,i}, \boldsymbol{S}_{t-1} = \boldsymbol{y}))_{s\times s}, \\ &\qquad j = 1, 2, \dots, m. \end{aligned}$$

In a similar fashion as Section 3, we can obtain the joint probability distribution of $(X_n, S_{n,1}, \dots, S_{n,m})$ by evaluating the $\boldsymbol{f}_n(x, \boldsymbol{y})$. The next theorem provides a method for the evaluation of the joint probability distribution of $(X_n, S_{n,1}, \dots, S_{n,m})$.

THEOREM 5.1. *The probability vectors $\boldsymbol{f}_t(x, \boldsymbol{y})$, $(t \geq 2)$ satisfy the recurrence relations*

$$\begin{aligned} \boldsymbol{f}_t(x, \boldsymbol{y}) &= \boldsymbol{f}_{t-1}(x, \boldsymbol{y})A_{t,0}(x, \boldsymbol{y}) + \sum_{j=1}^{m} \boldsymbol{f}_{t-1}(x, \boldsymbol{y} - \boldsymbol{e}_j)A_{t,j}(x, \boldsymbol{y} - \boldsymbol{e}_j) \\ &\quad + \boldsymbol{f}_{t-1}(x-1, \boldsymbol{y})B_{t,0}(x-1, \boldsymbol{y}) \\ &\quad + \sum_{j=1}^{m} \boldsymbol{f}_{t-1}(x-1, \boldsymbol{y} - \boldsymbol{e}_j)B_{t,j}(x-1, \boldsymbol{y} - \boldsymbol{e}_j) \\ &\quad + \boldsymbol{f}_{t-1}(x+1, \boldsymbol{y})C_{t,0}(x+1, \boldsymbol{y}) \\ &\quad + \sum_{j=1}^{m} \boldsymbol{f}_{t-1}(x+1, \boldsymbol{y} - \boldsymbol{e}_j)C_{t,j}(x+1, \boldsymbol{y} - \boldsymbol{e}_j), \\ &\quad \textit{for} \quad t \geq 2, \ 0 \leq x \leq \ell_t \quad \textit{and} \quad 0 \leq y_1, y_2, \dots, y_m \leq t - xn_0, \end{aligned} \tag{5.1}$$

with initial conditions

$$\begin{aligned} \boldsymbol{f}_1(0, \boldsymbol{y}) = (&\Pr(Y_1 = U_{0,0}, \boldsymbol{S}_1 = \boldsymbol{y}), \Pr(Y_1 = U_{0,1}, \boldsymbol{S}_1 = \boldsymbol{y}), \dots, \\ &\Pr(Y_1 = U_{0,s-1}, \boldsymbol{S}_1 = \boldsymbol{y})), \end{aligned}$$

for $y_1, y_2, \dots, y_m = 0, 1$. In addition, the joint probability distribution function of $(X_n, S_{n,1}, \dots, S_{n,m})$ is given by

$$\Pr(X_n = x, \boldsymbol{S}_n = \boldsymbol{y}) = \boldsymbol{f}_n(x, \boldsymbol{y})\mathbf{1}', \quad 0 \leq x \leq \ell_n, \ 0 \leq y_1, y_2, \dots, y_m \leq n - xn_0. \tag{5.2}$$

PROOF. Note that the event $\{Y_t = U_{x,j}, \boldsymbol{S}_t = \boldsymbol{y}\}$ implies the occurrence of one of the events

$$\bigcup_{j=0}^{s-1}\left(\{Y_{t-1} = U_{x,j}, \boldsymbol{S}_{t-1} = \boldsymbol{y}\} \cup \bigcup_{i=1}^{m}\{Y_{t-1} = U_{x,j}, \boldsymbol{S}_{t-1} = \boldsymbol{y} - \boldsymbol{e}_i\}\right)$$

or

$$\bigcup_{j=0}^{s-1}\left(\{Y_{t-1} = U_{x-1,j}, \boldsymbol{S}_{t-1} = \boldsymbol{y}\} \cup \bigcup_{i=1}^{m}\{Y_{t-1} = U_{x-1,j}, \boldsymbol{S}_{t-1} = \boldsymbol{y} - \boldsymbol{e}_i\}\right)$$

or

$$\bigcup_{j=0}^{s-1}\left(\{Y_{t-1} = U_{x+1,j}, \boldsymbol{S}_{t-1} = \boldsymbol{y}\} \cup \bigcup_{i=1}^{m}\{Y_{t-1} = U_{x+1,j}, \boldsymbol{S}_{t-1} = \boldsymbol{y} - \boldsymbol{e}_i\}\right).$$

The recurrences (5.1) are immediate consequences of the total probability theorem. It is easy to check the equation (5.2) from the following formula,

$$\Pr(X_n = x, \boldsymbol{S}_n = \boldsymbol{y}) = \Pr(Y_n \in \boldsymbol{U}_x, \boldsymbol{S}_n = \boldsymbol{y}) = \sum_{j=0}^{s-1} \Pr(Y_n = U_{x,j}, \boldsymbol{S}_n = \boldsymbol{y}). \qquad \square$$

Remark 5.1. In the special case where the matrices $A_{t,j}(x, \boldsymbol{y})$, $B_{t,j}(x, \boldsymbol{y})$ and $C_{t,j}(x, \boldsymbol{y})$ $(j = 0, 1, \dots, m)$ do not depend on $\boldsymbol{y}$, that is,

$$A_{t,j}(x, \boldsymbol{y}) = A_{t,j}(x), \qquad B_{t,j}(x, \boldsymbol{y}) = B_{t,j}(x), \qquad C_{t,j}(x, \boldsymbol{y}) = C_{t,j}(x), \qquad \text{for all } \boldsymbol{y},$$

then the matrices $\sum_{j=0}^{m} A_{t,j}(x)$, $\sum_{j=0}^{m} B_{t,j}(x)$ and $\sum_{j=0}^{m} C_{t,j}(x)$ are equal to the within state transition matrix $A_t(x)$, the between state transition matrix $B_t(x)$ and the return state transition matrix $C_t(x)$, respectively. Therefore, the matrix $\sum_{j=0}^{m}(A_{t,j}(x) + B_{t,j}(x) + C_{t,j}(x))$ is a stochastic matrix.

If $A_{t,j}(x, \boldsymbol{y}) = A_{t,j}$, $B_{t,j}(x, \boldsymbol{y}) = B_{t,j}$ and $C_{t,j}(x, \boldsymbol{y}) = C_{t,j}$ $(j = 0, 1, \dots, m)$ for all $x, \boldsymbol{y}$, then the joint pgf $\phi_n(u, \boldsymbol{v})$ of $(X_n, S_{n,1}, \dots, S_{n,m})$ defined as (2.1) can be expressed as a product in the following way.

THEOREM 5.2. *If $A_{t,j}(x, \boldsymbol{y}) = A_{t,j}$, $B_{t,j}(x, \boldsymbol{y}) = B_{t,j}$ and $C_{t,j}(x, \boldsymbol{y}) = C_{t,j}$ $(j = 0, 1, \dots, m)$ for all $x, \boldsymbol{y}$, then the joint pgf $\phi_n(u, \boldsymbol{v})$ of $(X_n, S_{n,1}, \dots, S_{n,m})$ can be expressed as*

$$(5.3) \quad \phi_n(u, \boldsymbol{v}) = \boldsymbol{a}(\boldsymbol{v}) \prod_{t=2}^{n} \left[A_{t,0} + uB_{t,0} + \frac{1}{u}C_{t,0} + \sum_{j=1}^{m} v_j \left(A_{t,j} + uB_{t,j} + \frac{1}{u}C_{t,j}\right)\right] \boldsymbol{1}',$$

where $\boldsymbol{a}(\boldsymbol{v})$ is given by (3.4).

PROOF. Introducing the vector generating functions

$$\boldsymbol{\phi}_t(u, \boldsymbol{v}) = \sum_{x=0}^{\ell_t} \sum_{y=0}^{t - xn_0} \boldsymbol{f}_t(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}},$$

we can write

$$\phi_t(u, \boldsymbol{v}) = \boldsymbol{\phi}_t(u, \boldsymbol{v})\mathbf{1}'.$$

By summing both sides of the equation (5.1) after multiplying by $u^x \boldsymbol{v}^{\boldsymbol{y}}$, $\boldsymbol{\phi}_t(u, \boldsymbol{v})$ can be expressed as

$$\begin{aligned}\boldsymbol{\phi}_t(u, \boldsymbol{v}) = &\sum_{x=0}^{\ell_t} \sum_{\boldsymbol{y}=\mathbf{0}}^{t-1-xn_0} \boldsymbol{f}_{t-1}(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} A_{t,0} + \sum_{j=1}^{m} v_j \sum_{x=0}^{\ell_t} \sum_{\boldsymbol{y}=\mathbf{0}}^{t-1-xn_0} \boldsymbol{f}_{t-1}(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} A_{t,j} \\ &+ u \sum_{x=0}^{\ell_t - 1} \sum_{\boldsymbol{y}=\mathbf{0}}^{t-1-xn_0} \boldsymbol{f}_{t-1}(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} B_{t,0} \\ &+ u \sum_{j=1}^{m} v_j \sum_{x=0}^{\ell_t - 1} \sum_{\boldsymbol{y}=\mathbf{0}}^{t-1-xn_0} \boldsymbol{f}_{t-1}(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} B_{t,j} \\ &+ \frac{1}{u} \sum_{x=0}^{\ell_t} \sum_{\boldsymbol{y}=\mathbf{0}}^{t-1-xn_0} \boldsymbol{f}_{t-1}(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} C_{t,0} \\ &+ \frac{1}{u} \sum_{j=1}^{m} v_j \sum_{x=0}^{\ell_t} \sum_{\boldsymbol{y}=\mathbf{0}}^{t-1-xn_0} \boldsymbol{f}_{t-1}(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} C_{t,j}.\end{aligned}$$

We consider the two possible cases $\ell_t = \ell_{t-1}$ and $\ell_t = \ell_{t-1} + 1$ separately. If $\ell_t = \ell_{t-1}$, we should note the identity

$$\boldsymbol{f}_{t-1}(\ell_{t-1}, \boldsymbol{y}) B_{t,j} = \mathbf{0}, \qquad \text{for} \quad j = 0, 1, \ldots, m.$$

If $\ell_t = \ell_{t-1} + 1$, we should note the identity

$$\boldsymbol{f}_{t-1}(\ell_t, \boldsymbol{y}) A_{t,j} = \mathbf{0}, \qquad \boldsymbol{f}_{t-1}(\ell_t, \boldsymbol{y}) C_{t,j} = \mathbf{0}, \qquad \text{for} \quad j = 0, 1, \ldots, m.$$

Hence, in both cases we have

$$\boldsymbol{\phi}_t(u, \boldsymbol{v}) = \boldsymbol{\phi}_{t-1}(u, \boldsymbol{v}) \left[A_{t,0} + uB_{t,0} + \frac{1}{u} C_{t,0} + \sum_{j=1}^{m} v_j \left(A_{t,j} + uB_{t,j} + \frac{1}{u} C_{t,j} \right) \right].$$

If we take into account that

$$\boldsymbol{\phi}_1(u, \boldsymbol{v}) = \sum_{x=0}^{\ell_1} \sum_{\boldsymbol{y}=\mathbf{0}}^{1} \boldsymbol{f}_1(x, \boldsymbol{y}) u^x \boldsymbol{v}^{\boldsymbol{y}} = \sum_{\boldsymbol{y}=\mathbf{0}}^{1} \boldsymbol{f}_1(0, \boldsymbol{y}) \boldsymbol{v}^{\boldsymbol{y}} = \boldsymbol{a}(\boldsymbol{v}),$$

the proof is completed. □

For the homogeneous case (i.e. $A_{t,j} = A_j$, $B_{t,j} = B_j$, $C_{t,j} = C_j$, $j = 0, 1, \ldots, m$), the double generating function of $(X_n, S_{n,1}, \ldots, S_{n,m})$ defined as (2.2) takes more compact form.

THEOREM 5.3. *If $A_{t,j}(x, \boldsymbol{y}) = A_j$, $B_{t,j}(x, \boldsymbol{y}) = B_j$ and $C_{t,j}(x, \boldsymbol{y}) = C_j$ ($j = 0, 1, \ldots, m$) for all $x, \boldsymbol{y}$ and $t \geq 1$, then the double generating function of $(X_n, S_{n,1}, \ldots,$*

$S_{n,m}$) is given by

$$\Phi(u, \boldsymbol{v}; w) = w\boldsymbol{a}(\boldsymbol{v})\left[I - w\left(A_0 + uB_0 + \frac{1}{u}C_0 + \sum_{j=1}^{m} v_j\left(A_j + uB_j + \frac{1}{u}C_j\right)\right)\right]^{-1}\mathbf{1}', \tag{5.4}$$

where $\boldsymbol{a}(\boldsymbol{v})$ is given by (3.4).

PROOF. From the equation (5.3) under the conditions that $A_{t,j}(x, \boldsymbol{y}) = A_j$, $B_{t,j}(x, \boldsymbol{y}) = B_j$ and $C_{t,j}(x, \boldsymbol{y}) = C_j$, $(j = 0, 1, \dots, m)$, the joint generating function of $\phi_n(u, \boldsymbol{v})$ can be expressed as

$$\phi_n(u, \boldsymbol{v}) = \boldsymbol{a}(\boldsymbol{v})\left[A_0 + uB_0 + \frac{1}{u}C_0 + \sum_{j=1}^{m} v_j\left(A_j + uB_j + \frac{1}{u}C_j\right)\right]^{n-1}\mathbf{1}'.$$

Making use of the formula

$$\sum_{n=0}^{\infty}\left[A_0 + uB_0 + \frac{1}{u}C_0 + \sum_{j=1}^{m} v_j\left(A_j + uB_j + \frac{1}{u}C_j\right)\right]^{n} w^n$$
$$= \left[I - w\left(A_0 + uB_0 + \frac{1}{u}C_0 + \sum_{j=1}^{m} v_j\left(A_j + uB_j + \frac{1}{u}C_j\right)\right)\right]^{-1},$$

we have the desired result. □

In a similar fashion as in the conclusion of Section 3, we can easily establish formulae for the double generating function of $(X_n, F_{n,0}, S_{n,1}, \dots, S_{n,m})$. Through $\phi_n(u, \boldsymbol{v})$, the joint pgf $\psi_n(u, v_0, \boldsymbol{v})$ of $(X_n, F_{n,0}, S_{n,1}, \dots, S_{n,m})$ can be expressed as

$$\begin{aligned}\psi_n(u, v_0, \boldsymbol{v}) &= E(u^{X_n} v_0^{F_{n,0}} v_1^{S_{n,1}} \cdots v_m^{S_{n,m}}) \\ &= E(u^{X_n} v_0^{n - S_{n,1} - S_{n,2} - \cdots - S_{n,m}} v_1^{S_{n,1}} \cdots v_m^{S_{n,m}}) \\ &= v_0^n \phi_n(u, v_1/v_0, \dots, v_m/v_0).\end{aligned}$$

Therefore the double generating function

$$\Psi(u, v_0, \boldsymbol{v}; w) = \sum_{n=1}^{\infty} \psi_n(u, v_0, \boldsymbol{v}) w^n$$

takes the form $\Psi(u, v_0, \boldsymbol{v}; w) = \Phi(u, v_1/v_0, \dots, v_m/v_0; wv_0)$. More specifically, in terms of (5.4), we have the following result.

THEOREM 5.4. *If $A_{t,j}(x, \boldsymbol{y}) = A_j$, $B_{t,j}(x, \boldsymbol{y}) = B_j$, $C_{t,j}(x, \boldsymbol{y}) = C_j$ $(j = 0, 1, \dots, m)$ for all $x, \boldsymbol{y}$ and $t \geq 1$, then the double generating function $\Psi(u, v_0, \boldsymbol{v}; w)$ of*

$(X_n, F_{n,0}, S_{n,1}, \ldots, S_{n,m})$ can be expressed as

$$(5.5)\qquad \Psi(u, v_0, \boldsymbol{v}; w) = wv_0\boldsymbol{a}(v_1/v_0, \ldots, v_m/v_0)\left[I - w\sum_{j=0}^{m} v_j\left(A_j + uB_j + \frac{1}{u}C_j\right)\right]^{-1}\mathbf{1}',$$

where $\boldsymbol{a}(\boldsymbol{v})$ is given by (3.4).

As by-products, we can derive some interesting formulae from the equation (5.5). The double generating functions of $(X_n, F_{n,0})$, $(X_n, S_{n,i})$ $i = 1, 2, \ldots, m$ are given by

$$\begin{aligned}
\sum_{n=1}^{\infty} E(u^{X_n} v_0^{F_{n,0}})w^n
&= \Psi(u, v_0, \boldsymbol{v}; w)\Big|_{v_1=v_2=\cdots=v_m=1} \\
&= wv_0\boldsymbol{a}(1/v_0, \ldots, 1/v_0) \\
&\quad \times \left[I - w\left(v_0\left(A_0 + uB_0 + \frac{1}{u}C_0\right) + \sum_{j=1}^{m}\left(A_j + uB_j + \frac{1}{u}C_j\right)\right)\right]^{-1}\mathbf{1}',
\end{aligned}$$

$$\begin{aligned}
\sum_{n=1}^{\infty} E(u^{X_n} v_i^{S_{n,i}})w^n
&= \Psi(u, v_0, \boldsymbol{v}; w)\Big|_{v_0=v_1=\cdots=v_{i-1}=1, v_{i+1}=\cdots=v_m=1} \\
&= w\boldsymbol{a}(\underbrace{1, \ldots, 1}_{i-1}, v_i, \underbrace{1, \ldots, 1}_{m-i}) \\
&\quad \times \left[I - w\left(v_i\left(A_i + uB_i + \frac{1}{u}C_i\right) + \sum_{j\neq i}^{m}\left(A_j + uB_j + \frac{1}{u}C_j\right)\right)\right]^{-1}\mathbf{1}', \\
&\qquad i = 1, 2, \ldots, m.
\end{aligned}$$

The double generating function of X_n is

$$\sum_{n=1}^{\infty} E(w^{X_n})u^n = \Psi(u, v_0, \boldsymbol{v}; w)\,|_{v_0=v_1=\cdots=v_m=1} = w\boldsymbol{a}(\mathbf{1})\left[I - w\left(A + uB + \frac{1}{u}C\right)\right]^{-1}\mathbf{1}',$$

where, $A = \sum_{j=0}^{m} A_j$, $B = \sum_{j=0}^{m} B_j$, $C = \sum_{j=0}^{m} C_j$ and $\boldsymbol{a}(\mathbf{1}) = \boldsymbol{a}(\boldsymbol{v})\,|_{v_1=\cdots=v_m=1}$ (for a direct proof of an analogous result see Han and Aki (1999)).

THEOREM 5.5. *If $A_{t,j}(x, \boldsymbol{y}) = A_j$, $B_{t,j}(x, \boldsymbol{y}) = B_j$, $C_{t,j}(x, \boldsymbol{y}) = C_j$ ($j = 0, 1, \ldots, m$) for all $x, \boldsymbol{y}$ and $t \geq 1$, then the expected value of X_n and its generating function are given by*

$$M_{X_n}(w) = \sum_{n=1}^{\infty} E(X_n)w^n = \frac{w^2}{1-w}\boldsymbol{a}(\mathbf{1})[I - w(A + B + C)]^{-1}(B - C)\mathbf{1}',$$

$$E(X_n) = \boldsymbol{a}(\mathbf{1}) \sum_{i=1}^{n-1} (A + B + C)^{i-1} (B - C)\mathbf{1}',$$

where, $\boldsymbol{a}(\mathbf{1}) = \boldsymbol{a}(\boldsymbol{v}) \mid_{v_1 = v_2 = \cdots = v_m = 1}$, $A = \sum_{j=0}^{m} A_j$, $B = \sum_{j=0}^{m} B_j$ *and* $C = \sum_{j=0}^{m} C_j$.

PROOF. Note that

$$\sum_{n=1}^{\infty} E(X_n) w^n = \frac{\partial}{\partial u} \Psi(u, v_0, \boldsymbol{v}; w) \Bigg|_{u = v_0 = v_1 = \cdots = v_m = 1}$$

and making use of

$$\frac{d}{dz}(U - zV)^{-1} = (U - zV)^{-1} V (U - zV)^{-1},$$

we have the first conclusion of the theorem. By expanding $M_{X_n}(w)$ in a power series of w, we have the expected value of X_n immediately. □

The expressions in Theorem 5.5 are easily shown to be consistent with similar expressions given in Han and Aki (1999). The formulae for the variance of X_n, the covariance between X_n and $F_{n,0}$ the covariances between X_n and $S_{n,i}$ $(i = 1, 2, \ldots, m)$ can be derived from the derivatives of the double generating function $\Psi(u, v_0, \boldsymbol{v}; w)$, however, their expressions are not very attractive.

Remark 5.2. When the number X_n of occurrences of $\mathcal{E}$ in n trials is an M.V.R., we can treat the joint distribution of the waiting time T_r until the r-th occurrence of the pattern $\mathcal{E}$, the numbers $S_{T_r,i}$ of outcomes "i" $(i = 1, 2, \ldots, m)$ and the number $F_{T_r,0}$ of outcomes "0" until T_r. However, we do not dwell on the matter, since the analysis of such case can be performed by combining the results in Section 4 with the results in Section 5.

Remark 5.3. Another important random variables are the numbers of an outcome "i" whose previous outcome is "j" $(i, j = 0, 1, \ldots, m)$ (i.e. the numbers of times that an outcome "i" follows an outcome "j"). For example, in DNA sequence, the probability structure of occurrences of four letters (A, T, G and C) is usually described by a first order Markov chain. Sometimes, a first order Markov model fits real data better than the independence model. It is important to assess whether the first order Markov dependence model describes reality better than independence model. The distribution of the numbers of times that an outcome "j" is followed by an outcome "i" $(j, i = A, T, G, C)$ is needed for the test for the first order Markov independence and provides more useful information for the modeling of the DNA sequence. Our methods and results in this paper can be extended to cover this case easily.

6. Applications

In this section, we will present some examples, which are closely related to practical problems, such as quality control, start-up demonstration test, reliability theory, ... etc. Assume that the transition matrices treated here are independent of x, $\boldsymbol{y}$ and t, that is, $A_{t,j}(x, \boldsymbol{y}) = A_j$, $B_{t,j}(x, \boldsymbol{y}) = B_j$, $C_{t,j}(x, \boldsymbol{y}) = C_j$ $(j = 0, 1, \ldots, m)$.

6.1 *Sooner waiting time problems*

Let $Z_1, Z_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables taking values in $\mathcal{A} = \{0, 1, \ldots, m\}$. Assume that

$$p_i = \Pr(Z_t = i), \qquad t \geq 1 \quad \text{and} \quad i = 0, 1, \ldots, m.$$

Let T_r $(r \geq 1)$ be the waiting time for r runs in total which are among "i"-runs of length k_i $(i = 1, 2, \ldots, m)$. Assume that counting of all runs are performed in the non-overlapping sense.

We have

$$A_0 + uB_0 + \sum_{i=1}^{m} v_i(A_i + uB_i)$$

$$= \left(\begin{array}{cccccccccccccc}
p_0 & p_1v_1 & 0 & \cdots & 0 & p_2v_2 & 0 & \cdots & 0 & \cdots & p_mv_m & 0 & \cdots & 0 \\
p_0 & 0 & p_1v_1 & \cdots & 0 & p_2v_2 & 0 & \cdots & 0 & \cdots & p_mv_m & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\
p_0 & 0 & 0 & \cdots & p_1v_1 & p_2v_2 & 0 & \cdots & 0 & \cdots & p_mv_m & 0 & \cdots & 0 \\
p_0 + up_1v_1 & 0 & 0 & \cdots & 0 & p_2v_2 & 0 & \cdots & 0 & \cdots & p_mv_m & 0 & \cdots & 0 \\
p_0 & p_1v_1 & 0 & \cdots & 0 & p_2v_2 & 0 & \cdots & 0 & \cdots & p_mv_m & 0 & \cdots & 0 \\
p_0 & p_1v_1 & 0 & \cdots & 0 & 0 & p_2v_2 & \cdots & 0 & \cdots & p_mv_m & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & p_mv_m & 0 & \cdots & 0 \\
p_0 & p_1v_1 & 0 & \cdots & 0 & 0 & 0 & \cdots & p_2v_2 & \cdots & p_mv_m & 0 & \cdots & 0 \\
p_0 + up_2v_2 & p_1v_1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots & p_mv_m & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\
p_0 & p_1v_1 & 0 & \cdots & 0 & p_2v_2 & 0 & \cdots & 0 & \cdots & p_mv_m & 0 & \cdots & 0 \\
p_0 & p_1v_1 & 0 & \cdots & 0 & p_2v_2 & 0 & \cdots & 0 & \cdots & 0 & p_mv_m & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & \ddots & \vdots \\
p_0 & p_1v_1 & 0 & \cdots & 0 & p_2v_2 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & p_mv_m \\
p_0 + up_mv_m & p_1v_1 & 0 & \cdots & 0 & p_2v_2 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0
\end{array} \right)_{s \times s},$$

where $s = \sum_{i=1}^{m} k_i - m + 1$. Since the initial condition

$$\boldsymbol{a}(\boldsymbol{v}) = (p_0, \underbrace{p_1v_1, 0, \ldots, 0}_{k_1-1}, \underbrace{p_2v_2, 0, \ldots, 0}_{k_2-1}, \ldots, \underbrace{p_mv_m, 0, \ldots, 0}_{k_m-1}),$$

and

$$\beta_{i,j} = \boldsymbol{e}_i B_j \boldsymbol{1}' = \begin{cases} p_1, & \text{if} \quad (i,j) = (k_1, 1), \\ p_2, & \text{if} \quad (i,j) = (k_1 + k_2 - 1, 1), \\ \vdots & \qquad \vdots \\ p_\ell, & \text{if} \quad (i,j) = (k_1 + k_2 + \cdots + k_\ell - \ell + 1, 1), \\ \vdots & \qquad \vdots \\ p_m, & \text{if} \quad (i,j) = (k_1 + k_2 + \cdots + k_m - m + 1, 1), \\ 0, & \text{otherwise}, \end{cases}$$

after some calculations, we obtain

$$H(w, \boldsymbol{v}; u) = \frac{uF_1(w, \boldsymbol{v})}{1 - p_0 w - F_0(w, \boldsymbol{v}) - uF_1(w, \boldsymbol{v})},$$

where,

$$F_0(w, \boldsymbol{v}) = \sum_{i=1}^{m} \frac{p_i w v_i - (p_i w v_i)^{k_i}}{1 - (p_i w v_i)^{k_i}},$$

$$F_1(w, \boldsymbol{v}) = \sum_{i=1}^{m} \frac{(p_i w v_i)^{k_i}(1 - p_i w v_i)}{1 - (p_i w v_i)^{k_i}}.$$

Furthermore, expanding $H(w, \boldsymbol{v}; u)$ in a power series of u, we obtain

$$H_r(w, \boldsymbol{v}) = \left(\frac{F_1(w, \boldsymbol{v})}{1 - p_0 w - F_0(w, \boldsymbol{v})}\right)^r.$$

In case of $r = 1$, we obtain the joint pgf of the sooner waiting time random variable T_1 and the numbers of successes $(S_{T_1,1}, \dots, S_{T_1,m})$ appeared at that time. Stefanov (2000) consider the waiting time problems for the first occurrence of a run of either $1'$s of length k_1 or $2'$s of length k_2 or $3'$s of length k_3 in a Markov chain with state space $\{1, 2, 3\}$ and also discussed other waiting time problems (see Stefanov and Pakes (1997, 1999)). Ebneshahrashoob and Sobel (1990), Feller (1968), Aki and Hirano (1993) and Aki *et al.* (1996)). Aki and Hirano (1994) studied the distribution of the number of failures and successes until the first occurrence of a success run of length k in some $\{0, 1\}$-valued random sequences (see Aki and Hirano (1995)).

As Koutras and Alexandrou (1997*b*) stated, it is important to consider the sooner waiting time problems, since their distributions play a key role in various applied areas of research (for example, quality control, start-up demonstration tests, learning criteria in psychology).

6.2 *The number of occurrences of success runs of exact length k*

Let $Z_0, Z_1, Z_2, \dots, Z_n$ be a time homogeneous $\{0, 1\}$-valued Markov chain with transition probabilities

$$p_{ij} = \Pr(Z_t = j \mid Z_{t-1} = i), \qquad t \geq 1 \quad \text{and} \quad i, j = 0, 1$$

and initial probabilities

$$p_j = \Pr(Z_0 = j), \qquad j = 0, 1,$$

(we say success and failure for the outcomes "1" and "0", respectively). We study the joint distribution of the number of occurrences of success runs of exact length k and the number of successes in the $Z_1, Z_2, \dots, Z_n$ (see Mood (1940), Han and Aki (1999) and Doi and Yamamoto (1998)).

We have

$$A_0 + uB_0 + \frac{1}{u}C_0 + v_1\left(A_1 + uB_1 + \frac{1}{u}C_1\right)$$

$$= \begin{pmatrix} p_{00} & p_{01}v_1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ p_{10} & 0 & p_{11}v_1 & 0 & \cdots & 0 & 0 & 0 \\ p_{10} & 0 & 0 & p_{11}v_1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ p_{10} & 0 & 0 & 0 & \cdots & p_{11}v_1 & 0 & 0 \\ p_{10} & 0 & 0 & 0 & \cdots & 0 & up_{11}v_1 & 0 \\ p_{10} & 0 & 0 & 0 & \cdots & 0 & 0 & p_{11}v_1/u \\ p_{10} & 0 & 0 & 0 & \cdots & 0 & 0 & p_{11}v_1 \end{pmatrix}_{(k+2)\times(k+2)}.$$

The initial condition is

$$\boldsymbol{a}(v_1) = (p_0p_{00} + p_1p_{10}, (p_0p_{01} + p_1p_{11})v_1, 0, \dots, 0) \in \mathcal{R}^{k+2}.$$

In order to obtain the double generating function of $(X_n, S_{n,1})$, we need to invert the matrix $I - w(A_0 + uB_0 + \frac{1}{u}C_0 + v_1(A_1 + uB_1 + \frac{1}{u}C_1))$. Straightforward calculations yields that

$$\Phi(u, v_1; w) = (p_0p_{00} + p_1p_{10})w\frac{P_0(u, v_1, w)}{P(u, v_1, w)} + (p_0p_{01} + p_1p_{11})wv_1\frac{P_1(u, v_1, w)}{P(u, v_1, w)}$$

where,

$$\begin{aligned} P_0(u, v_1, w) &= 1 + (p_{01} - p_{11})wv_1 - p_{01}wv_1(p_{11}wv_1)^{k-1}(1 - p_{11}wv_1)(1 - u), \\ P_1(u, v_1, w) &= 1 + (p_{10} - p_{00})w - (1 - u)(1 - p_{11}wv_1)(p_{11}wv_1)^{k-1}[1 + (p_{10} - p_{00})w], \\ P(u, v_1, w) &= 1 - (p_{00} + p_{11}v_1)w + p_{00}p_{11}w^2v_1 \\ &\quad - p_{10}p_{01}w^2v_1[1 - (1 - u)(1 - p_{11}wv_1)(p_{11}wv_1)^{k-1}]. \end{aligned}$$

6.3 *Waiting for the first scan*

Let $Z_1, Z_2, \dots$ be a sequence of i.i.d. trials taking values in $\mathcal{A} = \{0, 1\}$. Assume that

$$p_0 = \Pr(Z_t = 0), \qquad p_1 = \Pr(Z_t = 1), \qquad t \geq 1.$$

Let T_k^m denote the waiting time for the first occurrence of a scan, which is defined as

$$T_k^m = \min\left\{n : \sum_{j=\max(n-m+1,1)}^{n} Z_j \geq k\right\}.$$

The distribution of T_k^m will be referred to as "geometric distribution of order k/m" (see Balakrishnan and Koutras (2002) and Glaz *et al.* (2001)). In case of $k = m$, the corresponding distribution is geometric distribution of order k (see Philippou *et al.* (1983) and Koutras (1996*b*)). For example, in the case of $k = 3$, $m = 4$, we consider the joint distribution of $(T_3^4, S_{T_3^4,1})$. We have

$$A_0 + v_1 A_1 = \begin{pmatrix} p_0 & p_1v_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & p_0 & 0 & 0 & p_1v_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_0 & 0 & 0 & p_1v_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_0 & 0 & 0 & p_1v_1 & 0 & 0 & 0 \\ p_0 & p_1v_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_0 & 0 \\ 0 & 0 & p_0 & 0 & 0 & p_1v_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_0 \\ 0 & 0 & 0 & p_0 & 0 & 0 & p_1v_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_0 & 0 & 0 & p_1v_1 & 0 & 0 & 0 \end{pmatrix}_{11\times 11}.$$

Since the initial condition

$$\boldsymbol{a}(v_1) = (p_0, p_1v_1, 0, \dots, 0),$$

and

$$\beta_{i,j} = \boldsymbol{e}_i B_j \mathbf{1}' = \begin{cases} p_1, & \text{if } \ (i,j) = (6,1), (7,1), (9,1), \\ 0, & \text{otherwise,} \end{cases}$$

by a direct application of Corollary 4.1, we have

$$H_1(w, v_1) = \frac{(p_1wv_1)^3(1 + 2p_0w - p_0p_1w^2v_1 - p_0^2p_1w^3v_1)}{1 - p_0w - p_0p_1w^2v_1 - p_0^2p_1^2w^4v_1^2 + p_0^3p_1^3w^6v_1^3}.$$

Acknowledgements

We wish to thank the editor and the referees for careful reading of our paper and helpful suggestions which led improved results. Particularly, we would like to thank one of referees for his helpful comments on Subsection 6.1, which led to considerable improvements in the paper.

References

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, *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 consecutive k 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.

Aki, S., Balakrishnan, N. and Mohanty, S. G. (1996). Sooner and later waiting time problems and failure runs in higher order Markov dependent trials, *Annals of the Institute of Statistical Mathematics*, **48**, 773–787.

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

Biggins, J. D. and Cannings, C. (1987). Markov renewal processes, counters and repeated sequences in Markov chains, *Advances in Applied Probability*, **19**, 521–545.

Blom, G. and Thorburn, D. (1982). How many random digits are required until given sequences are obtained?, *Journal of Applied Probability*, **19**, 518–531.

Chadjiconstantinidis, S., Antzoulakos, D. L. and Koutras, M. V. (2000). Joint distributions of successes, failures and patterns in enumeration problems, *Advances in Applied Probability*, **32**, 866–884.

Doi, M. and Yamamoto, E. (1998). On the joint distribution of runs in a sequence of multi-state trials, *Statistics & Probability Letters*, **39**, 133–141.

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

Ewens, W. J. and Grant, G. R. (2001). *Statistical Methods in Bioinformatics: An Introduction*, Springer, New York.

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

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, *Journal of the American Statistical Association*, **89**, 1050–1058.

Fu, J. C. and Lou, W. Y. W. (2000). Joint distribution of rises and falls, *Annals of the Institute of Statistical Mathematics*, **52**, 415–425.

Glaz, J., Naus, J. and Wallenstein, S. (2001). *Scan Statistics*, Springer, New York.

Han, Q. and Aki, S. (1999). Joint distributions of runs in a sequence of multi-state trials, *Annals of the Institute of Statistical Mathematics*, **51**, 419–447.

Inoue, K. and Aki, S. (2002). Generalized waiting time problems associated with pattern in Polya's urn scheme, *Annals of the Institute of Statistical Mathematics*, **54**, 681–688.

Inoue, K. and Aki, S. (2003). 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. (1996*a*). On a Markov chain approach for the study of reliability structures, *Journal of Applied Probability*, **33**, 357–367.

Koutras, M. V. (1996*b*). On a waiting time distribution in a sequence of Bernoulli trials, *Annals of the Institute of Statistical Mathematics*, **48**, 789–806.

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

Koutras, M. V. and Alexandrou, V. A. (1995). Runs, scans and urn model distributions: A unified Markov chain approach, *Annals of the Institute of Statistical Mathematics*, **47**, 743–766.

Koutras, M. V. and Alexandrou, V. A. (1997*a*). Non-parametric randomness tests based on success runs of fixed length, *Statistics & Probability Letters*, **32**, 393–404.

Koutras, M. V. and Alexandrou, V. A. (1997*b*). Sooner waiting time problems in a sequence of trinary trials, *Journal of Applied Probability*, **34**, 593–609.

Mood, A. M. (1940). The distribution theory of runs, *Annals of Mathematical Statistics*, **11**, 367–392.

Philippou, A. N., Georghiou, C. and Philippou, G. N. (1983). A generalized geometric distribution and some of its properties, *Statistics & Probability Letters*, **1**, 171–175.

Robin, S. and Daudin, J. J. (1999). Exact distribution of word occurrences in a random sequence of letters, *Journal of Applied Probability*, **36**, 179–193.

Stefanov, V. T. (2000). On some waiting time problems, *Journal of Applied Probability*, **37**, 756–764.

Stefanov, V. T. and Pakes, A. G. (1997). Explicit distributional results in pattern formation, *The Annals of Applied Probability*, **7**, 666–678.

Stefanov, V. T. and Pakes, A. G. (1999). Explict distributional results in pattern formation II, *Australian & New Zealand Journal of Statistics*, **41**, 79–90.

Uchida, M. (1998). On generating functions of waiting time problems for sequence patterns of discrete random variables, *Annals of the Institute of Statistical Mathematics*, **50**, 655–671.