Ann. Inst. Statist. Math.
Vol. 51, No. 3, 419–447 (1999)

JOINT DISTRIBUTIONS OF RUNS IN A SEQUENCE OF MULTI-STATE TRIALS

QING HAN* AND SIGEO AKI

Department of Informatics and Mathematical Science, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama-cho, Toyonaka, Osaka 560-8531, Japan

(Received March 10, 1997; revised April 24, 1998)

Abstract. In this paper we introduce a Markov chain imbeddable vector of multinomial type and a Markov chain imbeddable variable of returnable type and discuss some of their properties. These concepts are extensions of the Markov chain imbeddable random variable of binomial type which was introduced and developed by Koutras and Alexandrou (1995, *Ann. Inst. Statist. Math.*, **47**, 743–766). By using the results, we obtain the distributions and the probability generating functions of numbers of occurrences of runs of a specified length based on four different ways of counting in a sequence of multi-state trials. Our results also yield the distribution of the waiting time problems.

Key words and phrases: Run statistics, distributions of order k, probability generating function, Markov chain imbedding, multi-state trials, waiting time problem.

1. Introduction

In recent years exact discrete distribution theory related to run statistics has been developed. There are various definitions of run. The four most important and frequently used success runs are:

(a) $E_{n,k}$, the number of success runs of size exactly k until the n-th trial (Mood (1940));

(b) $N_{n,k}$, the number of nonoverlapping consecutive k successes until the n-th trial (Feller (1968));

(c) $M_{n,k}$, the number of overlapping consecutive k successes until the n-th trial (Ling (1988));

(d) $G_{n,k}$, the number of success runs of size greater than or equal to k until the n-th trial.

The reliability of consecutive-k-out-of-n: F systems is closely related to these run statistics (Aki (1985), Hirano (1986) and Philippou (1986)). The probability

* Now at Department of Statistics, East China Normal University, Shanghai, 200062, China.

generating functions (pgfs) and the probability functions (pfs) of these run statistics have been studied by many authors in many situations. Hirano (1986) and Philippou and Makri (1986) derived the pf of the $N_{n,k}$ in independent case. Ling (1988, 1989) studied the pf related to the $M_{n,k}$ in independent trials. Hirano *et al.* (1991) obtained the pgfs and the pfs of $M_{n,k}$ and $N_{n,k}$ in independent case. Aki and Hirano (1993) and Hirano and Aki (1993) obtained the pgfs and the pfs of $N_{n,k}$, $M_{n,k}$ and $G_{n,k}$ in a two-state Markov chain. Aki (1985) and Aki and Hirano (1988) discussed the pgfs and the pfs of some distributions related to runs in a binary sequence of order k. Schwager (1983) discussed the runs in $v \geq 2$ possible outcomes at each trial. Aki (1992) studied some waiting time problems in $\{0, 1, 2, \ldots\}$-valued random sequences. Uchida and Aki (1995) obtained the recurrence relations of the pgfs of the sooner and later waiting time problems in a two-state Markov chain. Ling and Low (1993) provided formulae for the waiting time until the r-th occurrence of a run of indentical symbols. Philippou *et al.* (1988) derived a multivariate negative binomial distribution of order k, by mean of an urn scheme.

Recently, Fu and Koutras (1994) studied the distributions of the most common run statistics ($E_{n,k}$, $N_{n,k}$, $M_{n,k}$, and $G_{n,k}$) based on a finite Markov chain imbedding technique. Because each of the transition probability matrices of the most common run statistics ($N_{n,k}$, $M_{n,k}$ and $G_{n,k}$) can be viewed as a bidiagonal matrix with non-zero blocks appearing only on the main diagonal and on the diagonal next to it, Koutras and Alexandrou (1995) introduced a Markov chain imbeddable variable of Binomial type (M.V.B.) and derived the recurrence relations of probability vectors. Using these relations, they obtained the distributions of the $N_{n,k}$, $M_{n,k}$ and $G_{n,k}$ run statistics and scan statistics. Fu (1996) introduced a "forward and backward principle" for the finite Markov chain imbedding to study the exact and joint distributions of the runs and patterns in a sequence of multi-state trials. Koutras and Alexandrou (1997) investigated the sooner waiting time distributions in both linear and circular problems by using the Markov chain imbedding method.

Fu and Koutras' approach provides a unified procedure for the evaluation of the distributions of runs, scans and patterns. A disadvantage of this approach is that the dimension of the transition probability matrices increase with the number n of trials. The evaluation is difficult at large sequences of trials. Koutras and Alexandrou improved the Markov chain approach for Bernoulli trials. Because the dimension of the probability vectors is independent of the number n of trials, the evaluation of the distribution of M.V.B. is easily performed recursively at large sequences of trials. However, we do not know whether the $E_{n,k}$ is M.V.B., since it is difficult to create the proper Markov chain which does not move backwards.

In this paper, we extend the concept of the M.V.B. A Markov chain imbeddable vector of multinomial type is introduced to study the joint distributions of the most common run statistics (E, N, M, and G) in a sequence of multi-state trials. The evaluation of the joint distributions of run statistics is easily performed recursively for multi-state trials. Further, we introduce the Markov chain imbeddable variable of returnable type, which enables the evaluation of the distribution of $E_{n,k}$ in a sequence of Bernoulli trials.

In Section 2, we introduce the Markov chain imbeddable vector of multinomial type, and give some basic results. In Section 3, we obtain the joint distributions for $N_{n,k}$, $M_{n,k}$ and $G_{n,k}$ in a sequence of multi-state trials, by applying the results of Section 2. A numerical example is given in Section 3. In Section 4, we introduce the Markov chain imbeddable variable (vector) of returnable type. Using the results of Section 4, we give the distribution for $E_{n,k}$ in a sequence of Bernoulli trials and the joint distribution of $E_{n,k}$ in a sequence of multi-state trials in Section 5. A numerical example is also given. In Section 6, we discuss the waiting time problems.

2. Markov chain imbeddable vector of multinomial type

In this section, we consider the run statistics of both identical and nonidentical independent sequence of the multi-state trials. Let $Y_1, Y_2, \dots$ be independent trials with $(m+1)$ possible outcomes "0", "1", ..., "m" in each trial. i.e. $Y_1, Y_2, \dots$ is a sequence of $\{0, 1, \dots, m\}$-valued random variable. Assume that $\Pr(Y_t = i) = p_t^{(i)}$, $i = 0, 1, \dots, m$ for $t = 1, \dots,$ and $(\sum_{i=0}^{m} p_t^{(i)} = 1)$, $(t = 1, 2, \dots)$.

We introduce the following notations:

(1) $N_{n,k_i}^{(i)}$ is the number of nonoverlapping "i"-runs of length k_i $(i = 0, 1, \dots, m)$;

(2) $M_{n,k_i}^{(i)}$ is the number of overlapping "i"-runs of length k_i $(i = 0, 1, \dots, m)$;

(3) $G_{n,k_i}^{(i)}$ is the number of "i"-runs of length greater than or equal to k_i, $(i = 0, 1, \dots, m)$;

(4) $E_{n,k_i}^{(i)}$ is the number of "i"-runs of size exactly k_i $(i = 0, 1, \dots, m)$, until the n-th trial.

Example. Suppose we have a sequence of trials with possible outcomes "0", "1", "2". If the outcomes are "11220222011100111122", then $n = 20$, $m = 2$ and we have $N_{20,2}^{(0)} = 1$, $N_{20,2}^{(1)} = 4$, $N_{20,2}^{(2)} = 3$, $M_{20,2}^{(0)} = 1$, $M_{20,2}^{(1)} = 6$, $M_{20,2}^{(2)} = 4$, $G_{20,2}^{(0)} = 1$, $G_{20,2}^{(1)} = 3$, $G_{20,2}^{(2)} = 3$, $E_{20,2}^{(0)} = 1$, $E_{20,2}^{(1)} = 1$, $E_{20,2}^{(2)} = 2$ and $N_{20,3}^{(0)} = 0$, $N_{20,3}^{(1)} = 2$, $N_{20,3}^{(2)} = 1$, ...

Clearly, when $m = 1$, it becomes the Bernoulli trials case, which was discussed by Fu and Koutras (1994) and Koutras and Alexandrou (1995).

For convenience, we use the random variable $X_n^{(i)}$ to represent "i"-run statistics $E_{n,k_i}^{(i)}$, $N_{n,k_i}^{(i)}$, $M_{n,k_i}^{(i)}$ and $G_{n,k_i}^{(i)}$.

Let $\boldsymbol{X}_n = (X_n^{(0)}, X_n^{(1)}, \dots, X_n^{(m)})$, ($X_n^{(i)}$ is a non-negative integer valued random variable). $\boldsymbol{x} = (x^{(0)}, x^{(1)}, \dots, x^{(m)})$ is a $\boldsymbol{X}_n$'s realization.

Let $l_n^{(i)} = \max\{x^{(i)} : \Pr(X_n^{(i)} = x^{(i)}) > 0\}$, $(i = 0, 1, \dots, m)$, and $\boldsymbol{l}_n = (l_n^{(0)}, l_n^{(1)}, \dots, l_n^{(m)})$. So, $\Pr(\boldsymbol{X}_n = \boldsymbol{x}) = 0$, for all $\boldsymbol{x} \notin \{\boldsymbol{x} : \boldsymbol{0} \le \boldsymbol{x} \le \boldsymbol{l}_n\}$.

Let $\boldsymbol{e}_k = (0, 0, \dots, 0, 1, 0, \dots, 0)_{1 \times (m+1)}$, (the $(k+1)$-th element is 1, and other elements are all 0), $(k = 0, 1, \dots, m)$; $\boldsymbol{1} = (1, 1, \dots, 1)$, ($(m+1)$ elements are all 1).

DEFINITION 2.1. The random vector $\boldsymbol{X}_n$ is called a Markov chain imbeddable vector of multinomial type (M.V.M.), if

(1) there exists a Markov chain $\{Z_t, t \geq 0\}$ defined on a state space Ω,
(2) there exists a partition $\{\boldsymbol{U_x} : \boldsymbol{x} \geq \boldsymbol{0}\}$ on the state space Ω,
(3) for every $\boldsymbol{x}$,

$$\Pr(\boldsymbol{X}_n = \boldsymbol{x}) = \Pr(Z_n \in \boldsymbol{U_x})$$

and (4)

$$\Pr(Z_t \in \boldsymbol{U}_{\boldsymbol{x}+\boldsymbol{x}^\star} \mid Z_{t-1} \in \boldsymbol{U_x}) = 0, \quad \text{if} \quad \boldsymbol{x}^\star \neq \boldsymbol{0}, \quad \text{or} \quad \boldsymbol{x}^\star \neq e_k, \qquad (k = 0, 1, \dots, m).$$

From (4), follows that the Markov chain cannot move backwards or jump directly to a higher state without visiting one of $(m+1)$ neighboring states $\boldsymbol{U}_{\boldsymbol{x}+e_k}$, $(k = 0, 1, \dots, m)$.

Without loss of generality, we assume that the sets $\boldsymbol{U_x}$ have the common cardinality $s = |\boldsymbol{U_x}|$ for every $\boldsymbol{x}$, so we denote $\boldsymbol{U_x} = \{U_{\boldsymbol{x},1}, U_{\boldsymbol{x},2}, \dots, U_{\boldsymbol{x},s}\}$.

We introduce $s \times s$ transition probability matrices

$$(2.1) \qquad A_t(\boldsymbol{x}) = (\Pr(Z_t = U_{\boldsymbol{x},j} \mid Z_{t-1} = U_{\boldsymbol{x},i}))_{s\times s},$$

$$(2.2) \qquad B_t^{(k)}(\boldsymbol{x}) = (\Pr(Z_t = U_{\boldsymbol{x}+e_k,j} \mid Z_{t-1} = U_{\boldsymbol{x},i}))_{s\times s}, \qquad (k = 0, 1, \dots, m).$$

Clearly, (1) the entries of $A_t(\boldsymbol{x})$ control the within state one-step transitions, (2) the entries of $B_t^{(k)}(\boldsymbol{x})$ control the between states $\boldsymbol{U_x}$ and $\boldsymbol{U}_{\boldsymbol{x}+e_k}$ one-step transitions.

We introduce the probability vectors of the t-th step Z_t of the Markov chain

$$\boldsymbol{f}_t(\boldsymbol{x}) = (\Pr(Z_t = U_{\boldsymbol{x},1}), \dots, \Pr(Z_t = U_{\boldsymbol{x},s})), \qquad \boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{l}_t, \qquad (t = 0, 1, \dots, n),$$

where $\boldsymbol{l}_t$ is $\boldsymbol{l}_t = (l_t^{(0)}, l_t^{(1)}, \dots, l_t^{(m)})$ and $l_t^{(i)}$ $(i = 0, 1, \dots, m)$ is the largest the $(i+1)$-th element in subscripts $\boldsymbol{x}$ of partitions $\boldsymbol{U_x}$ which include accessible state of the t-th step Z_t of the Markov chain, i.e.

$$l_t^{(i)} = \max\{x^{(i)} : \Pr(Z_t \in \boldsymbol{U}_{(\dots,x^{(i)},\dots)}) > 0\}, \qquad (i = 0, 1, \dots, m).$$

When $t = n$, it is $\boldsymbol{l}_n$ (from (3) of Definition 2.1). Let $\boldsymbol{\pi_x}$ be the initial probabilities of the Markov chain $\{Z_t, t \geq 0\}$, i.e. $\boldsymbol{\pi_x} = \boldsymbol{f}_0(\boldsymbol{x})$, $\boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{l}_0$. In the applications for run statistics, we have usually $\boldsymbol{l}_0 = \boldsymbol{0}$. Then we have

THEOREM 2.1. *The double sequence of probability vectors* $\boldsymbol{f}_t(\boldsymbol{x})$, $\boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{l}_t$, $t = 0, 1, \dots, n$, *satisfies*

$$(2.3) \quad \boldsymbol{f}_t(\boldsymbol{x}) = \boldsymbol{f}_{t-1}(\boldsymbol{x})A_t(\boldsymbol{x}) + \sum_{k=0}^{m} \boldsymbol{f}_{t-1}(\boldsymbol{x} - e_k)B_t^{(k)}(\boldsymbol{x} - e_k)I(\boldsymbol{x} - e_k \geq \boldsymbol{0}), \qquad (\boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{l}_t), \qquad (t = 1, \dots, n),$$

where

$$I(P) = \begin{cases} 1 & \text{if} \quad P \text{ is true} \\ 0 & \text{other} \end{cases}.$$

If X_n is a M.V.M., then the probability distribution function is given by

$$\Pr(\boldsymbol{X}_n = \boldsymbol{x}) = \boldsymbol{f}_n(\boldsymbol{x})\mathbf{1}'. \tag{2.4}$$

PROOF. The recurrence relations are immediate consequences of Chapman-Kolmogorov equations, Definition 2.1 and the form of the matrices $A_t(\boldsymbol{x})$, $B_t^{(k)}(\boldsymbol{x})$, $(k = 0, 1, \ldots, m)$. And

$$\Pr(\boldsymbol{X}_n = \boldsymbol{x}) = \Pr(Z_n \in \boldsymbol{U}_{\boldsymbol{x}}) = \sum_{i=1}^{s} \Pr(Z_n \in U_{\boldsymbol{x},i}) = \boldsymbol{f}_n(\boldsymbol{x})\mathbf{1}'.$$

This completes the proof.

The use of the nomenclature "multinomial type" is justified by the apparent similarity of formula (2.3) to the following relations

$$m(t; p_0, \ldots, p_m; x_0, \ldots, x_m) = \sum_{i=0}^{m} m(t-1; p_0, \ldots, p_m; x_0, \ldots, x_i - 1, \ldots, x_m) \cdot p_i,$$

where $m(n; p_0, \ldots, p_m; x_0, \ldots, x_m) = \frac{n!}{x_0! \cdots x_m!} p_0^{x_0} \cdots p_m^{x_m}$ is the pf of the multinomial distribution.

In most of the applications, the matrices $A_t(\boldsymbol{x})$, $B_t^{(k)}(\boldsymbol{x})$, $(k = 0, 1, \ldots, m)$ do not depend on $\boldsymbol{x}$. In the following of this section, we discuss this special case.

We consider the probability generating function of $\boldsymbol{X}_n$

$$\varphi_n(z_0, \ldots, z_m) = \sum_{\mathbf{0} \le \boldsymbol{x} \le \boldsymbol{l}_n} \Pr(\boldsymbol{X}_n = \boldsymbol{x}) z_0^{x_0} z_1^{x_1} \cdots z_m^{x_m}.$$

Let $\boldsymbol{z} = (z_0, \ldots, z_m)$, and let $\boldsymbol{z}^{\boldsymbol{x}} = z_0^{x_0} \cdots z_m^{x_m}$.

We define $\boldsymbol{\varphi}_t(\boldsymbol{z}) = \sum_{\mathbf{0} \le \boldsymbol{x} \le \boldsymbol{l}_t} \boldsymbol{f}_t(\boldsymbol{x}) \boldsymbol{z}^{\boldsymbol{x}}$ for the Markov chain, then $\varphi_n(\boldsymbol{z}) = \sum_{\mathbf{0} \le \boldsymbol{x} \le \boldsymbol{l}_n} \boldsymbol{f}_n(\boldsymbol{x}) \mathbf{1}' \boldsymbol{z}^{\boldsymbol{x}} = \boldsymbol{\varphi}_n(\boldsymbol{z}) \mathbf{1}'$.

We introduce $\chi_{(y)}^{(k)} = \{\boldsymbol{x} : x^{(k)} = y; 0 \le x^{(j)} \le l_t^{(j)}, j = 0, 1, \ldots, k-1, k+1, \ldots, m\}$, where $0 \le y \le l_t^{(k)}$ and $k = 0, 1, \ldots, m$.

THEOREM 2.2. *If $A_t(\boldsymbol{x})$, $B_t^{(k)}(\boldsymbol{x})$ do not depend on $\boldsymbol{x}$, and $A_t(\boldsymbol{x}) = A_t$, $B_t^{(k)}(\boldsymbol{x}) = B_t^{(k)}$, $(k = 0, 1, \ldots, m)$ for all $\boldsymbol{x}$, we have*

$$\boldsymbol{\varphi}_t(\boldsymbol{z}) = \boldsymbol{\varphi}_{t-1}(\boldsymbol{z}) \left(A_t + \sum_{k=0}^{m} z_k B_t^{(k)} \right). \tag{2.5}$$

PROOF. We have

$$\begin{aligned}
\varphi_t(z) &= \sum_{0\le x\le l_t} f_t(x)z^x \\
&= \sum_{0\le x\le l_t}\left[f_{t-1}(x)A_t + \sum_{k=0}^{m} f_{t-1}(x-e_k)B_t^{(k)}I(x-e_k\ge 0)\right]z^x \\
&= \sum_{0\le x\le l_t} f_{t-1}(x)A_t z^x + \sum_{k=0}^{m}\left[\sum_{0\le x\le l_t} f_{t-1}(x-e_k)B_t^{(k)}I(x-e_k\ge 0)z^x\right] \\
&= \left[\sum_{0\le x\le l_t} f_{t-1}(x)z^x\right]A_t \\
&\quad + \sum_{k=0}^{m}\left[z_k\left[\sum_{0\le x\le l_t} f_{t-1}(x-e_k)I(x-e_k\ge 0)z^{x-e_k}\right]B_t^{(k)}\right] \\
&= \left[\sum_{0\le x\le l_t} f_{t-1}(x)z^x\right]A_t + \sum_{k=0}^{m}\left[z_k\left[\sum_{0\le x\le l_t-e_k} f_{t-1}(x)z^x\right]B_t^{(k)}\right] \\
&= \left[\sum_{0\le x\le l_t} f_{t-1}(x)z^x\right]A_t \\
&\quad + \sum_{k=0}^{m}\left[z_k\left[\sum_{0\le x\le l_t} f_{t-1}(x)z^x - \sum_{x\in\mathcal{X}^{(k)}_{(l_t^{(k)})}} f_{t-1}(x)z^x\right]B_t^{(k)}\right].
\end{aligned}$$

Note that $l_t - l_{t-1}$ is $\mathbf{0}$ or a finite sum of different e_k's, $(k = 0, 1, \ldots, m)$. If $l_t - l_{t-1} = \mathbf{0}$, we have

$$\begin{aligned}
\varphi_t(z) &= \varphi_{t-1}(z)A_t + \sum_{k=0}^{m}\left[z_k\left[\sum_{0\le x\le l_{t-1}} f_{t-1}(x)z^x - \sum_{x\in\mathcal{X}^{(k)}_{(l_{t-1}^{(k)})}} f_{t-1}(x)z^x\right]B_t^{(k)}\right] \\
&= \varphi_{t-1}(z)\left[A_t + \sum_{k=0}^{m} z_k B_t^{(k)}\right] - \sum_{k=0}^{m}\left[z_k\left[\sum_{x\in\mathcal{X}^{(k)}_{(l_{t-1}^{(k)})}} f_{t-1}(x)z^x\right]B_t^{(k)}\right].
\end{aligned}$$

Setting $z = \mathbf{1}$ and multiplying both sides by $\mathbf{1}'$, because $\varphi_t(\mathbf{1})\mathbf{1}' = \varphi_t(\mathbf{1}) = \sum_{0\le x\le l_t}\Pr(X_t = x) = 1$ and $(A_t + \sum_{k=0}^{m} B_t^{(k)})\mathbf{1}' = \mathbf{1}'$, and $\varphi_{t-1}(\mathbf{1})\mathbf{1}' = 1$, we obtain

$$\sum_{k=0}^{m}\left[\sum_{\boldsymbol{x}\in\mathcal{X}^{(k)}_{(l^{(k)}_{t-1})}} \boldsymbol{f}_{t-1}(\boldsymbol{x})B^{(k)}_t\right]\mathbf{1}'=0.$$

Noting that each term of the summation is nonnegative, we get

$$\sum_{\boldsymbol{x}\in\mathcal{X}^{(k)}_{(l^{(k)}_{t-1})}} \boldsymbol{f}_{t-1}(\boldsymbol{x})B^{(k)}_t=\mathbf{0}\qquad (k=0,\dots,m).$$

Hence if $l^{(k)}_t=l^{(k)}_{t-1}$, $(k=0,1,\dots,m)$, we have

$$\boldsymbol{f}_{t-1}(\boldsymbol{x})B^{(k)}_t=\mathbf{0},\qquad \text{if}\quad \boldsymbol{x}\in\mathcal{X}^{(k)}_{(l^{(k)}_{t-1})}. \tag{2.6}$$

Generally, if $\boldsymbol{l}_t=\boldsymbol{l}_{t-1}+\sum_{j=0}^{r}\boldsymbol{e}_{i_j}$, we have

$$\begin{aligned}
\varphi_t(\boldsymbol{z}) &= \left[\sum_{0\le\boldsymbol{x}\le\boldsymbol{l}_{t-1}+\sum_{j=0}^{r}\boldsymbol{e}_{i_j}} \boldsymbol{f}_{t-1}(\boldsymbol{x})\boldsymbol{z}^{\boldsymbol{x}}\right]A_t \\
&\quad+\sum_{k=0}^{m} z_k\left[\sum_{0\le\boldsymbol{x}\le\boldsymbol{l}_{t-1}+\sum_{j=0}^{r}\boldsymbol{e}_{i_j}-\boldsymbol{e}_k} \boldsymbol{f}_{t-1}(\boldsymbol{x})\boldsymbol{z}^{\boldsymbol{x}}\right]B^{(k)}_t \\
&= \sum_{0\le\boldsymbol{x}\le\boldsymbol{l}_{t-1}} \boldsymbol{f}_{t-1}(\boldsymbol{x})\boldsymbol{z}^{\boldsymbol{x}}A_t+\sum_{j=0}^{r}\left[\sum_{\boldsymbol{x}\in\mathcal{X}^{(i_j)}_{(l^{(i_j)}_{t-1}+1)}} \boldsymbol{f}_{t-1}(\boldsymbol{x})\boldsymbol{z}^{\boldsymbol{x}}A_t\right] \\
&\quad+\left(\sum_{s=0}^{r}+\sum_{s=r+1}^{m}\right) z_{i_s}\left[\sum_{0\le\boldsymbol{x}\le\boldsymbol{l}_{t-1}+\sum_{j=0}^{r}\boldsymbol{e}_{i_j}-\boldsymbol{e}_{i_s}} \boldsymbol{f}_{t-1}(\boldsymbol{x})\boldsymbol{z}^{\boldsymbol{x}}\right]B^{(i_s)}_t \\
&= \varphi_{t-1}(\boldsymbol{z})A_t+\sum_{j=0}^{r}\left[\sum_{\boldsymbol{x}\in\mathcal{X}^{(i_j)}_{(l^{(i_j)}_{t})}} \boldsymbol{f}_{t-1}(\boldsymbol{x})\boldsymbol{z}^{\boldsymbol{x}}\right]A_t \\
&\quad+\sum_{s=0}^{r} z_{i_s}\left[\sum_{0\le\boldsymbol{x}\le\boldsymbol{l}_{t-1}+\sum_{j=0,j\ne s}^{r}\boldsymbol{e}_{i_j}} \boldsymbol{f}_{t-1}(\boldsymbol{x})\boldsymbol{z}^{\boldsymbol{x}}\right]B^{(i_s)}_t \\
&\quad+\sum_{s=r+1}^{m} z_{i_s}\left[\sum_{0\le\boldsymbol{x}\le\boldsymbol{l}_{t-1}-\boldsymbol{e}_{i_s}+\sum_{j=0}^{r}\boldsymbol{e}_{i_j}} \boldsymbol{f}_{t-1}(\boldsymbol{x})\boldsymbol{z}^{\boldsymbol{x}}\right]B^{(i_s)}_t
\end{aligned}$$

$$
\begin{aligned}
&= \varphi_{t-1}(z)A_t + \sum_{j=0}^{r}\left[\sum_{x\in\mathcal{X}^{(i_j)}_{(l_t^{(i_j)})}} f_{t-1}(x)z^x A_t\right] \\
&+ \sum_{s=0}^{r} z_{i_s}\left[\sum_{0\le x\le l_{t-1}} f_{t-1}(x)z^x\right] B_t^{(i_s)} \\
&+ \sum_{s=0}^{r} z_{i_s}\left[\sum_{j=0,j\neq s}^{r}\ \sum_{x\in\mathcal{X}^{(i_j)}_{(l_{t-1}^{(i_j)}+1)}} f_{t-1}(x)z^x\right] B_t^{(i_s)} \\
&+ \sum_{s=r+1}^{m} z_{i_s}\left[\sum_{0\le x\le l_{t-1}-e_{i_s}} f_{t-1}(x)z^x\right] B_t^{(i_s)} \\
&+ \sum_{s=r+1}^{m} z_{i_s}\left[\sum_{j=0}^{r}\ \sum_{x\in\mathcal{X}^{(i_j)}_{(l_{t-1}^{(i_j)}+1)}} f_{t-1}(x)z^x\right] B_t^{(i_s)} \\
&= \varphi_{t-1}(z)A_t + \sum_{j=0}^{r}\left[\sum_{x\in\mathcal{X}^{(i_j)}_{(l_t^{(i_j)})}} f_{t-1}(x)z^x A_t\right] + \sum_{s=0}^{r} z_{i_s}\varphi_{t-1}(z)B_t^{(i_s)} \\
&+ \sum_{s=0}^{r} z_{i_s}\left[\sum_{j=0,j\neq s}^{r}\ \sum_{x\in\mathcal{X}^{(i_j)}_{(l_t^{(i_j)})}} f_{t-1}(x)z^x\right] B_t^{(i_s)} \\
&+ \sum_{s=r+1}^{m} z_{i_s}\left[\sum_{0\le x\le l_{t-1}} f_{t-1}(x)z^x\right] B_t^{(i_s)} \\
&- \sum_{s=r+1}^{m} z_{i_s}\left[\sum_{x\in\mathcal{X}^{(i_s)}_{(l_{t-1}^{(i_s)})}} f_{t-1}(x)z^x\right] B_t^{(i_s)} \\
&+ \sum_{s=r+1}^{m} z_{i_s}\left[\sum_{j=0}^{r}\ \sum_{x\in\mathcal{X}^{(i_j)}_{(l_t^{(i_j)})}} f_{t-1}(x)z^x\right] B_t^{(i_s)}.
\end{aligned}
$$

From formula (2.6), $-\sum_{s=r+1}^{m}[z_{i_s}\sum_{x\in\mathcal{X}^{(i_s)}_{(l_{t-1}^{(i_s)})}} f_{t-1}(x)z^x B_t^{(i_s)}] = \mathbf{0}$.

Then by setting $\boldsymbol{z} = \mathbf{1}$, and multiplying by $\mathbf{1}'$, we have

$$\sum_{j=0}^{r}\left[\sum_{\boldsymbol{x}\in\mathcal{X}^{(i_j)}_{(l_t^{(i_j)})}} \boldsymbol{f}_{t-1}(\boldsymbol{x})A_t\right] + \sum_{s=0}^{r}\left[\sum_{j=0,j\neq s}^{r}\sum_{\boldsymbol{x}\in\mathcal{X}^{(i_j)}_{(l_t^{(i_j)})}} \boldsymbol{f}_{t-1}(\boldsymbol{x})B_t^{(i_s)}\right]$$
$$+\sum_{s=r+1}^{m}\left[\sum_{j=0}^{r}\sum_{\boldsymbol{x}\in\mathcal{X}^{(i_j)}_{(l_t^{(i_j)})}} \boldsymbol{f}_{t-1}(\boldsymbol{x})B_t^{(i_s)}\right] = \mathbf{0},$$

and every term is equal to $\mathbf{0}$ (because every term is nonnegative).

Hence we have

$$\boldsymbol{\varphi}_t(\boldsymbol{z}) = \boldsymbol{\varphi}_{t-1}(\boldsymbol{z})\left(A_t + \sum_{k=0}^{m} z_k B_t^{(k)}\right).$$

This completes the proof.

From this theorem, we have

$$\varphi_n(\boldsymbol{z}) = \boldsymbol{\varphi}_0(\boldsymbol{z})\prod_{t=1}^{n}\left(A_t + \sum_{k=0}^{m} z_k B_t^{(k)}\right)\mathbf{1}', \tag{2.7}$$

where

$$\boldsymbol{\varphi}_0(\boldsymbol{z}) = \sum_{\mathbf{0}\le\boldsymbol{x}\le\boldsymbol{l}_0} \boldsymbol{f}_0(\boldsymbol{x})\boldsymbol{z}^{\boldsymbol{x}} = \sum_{\mathbf{0}\le\boldsymbol{x}\le\boldsymbol{l}_0} \boldsymbol{\pi}_{\boldsymbol{x}}\boldsymbol{z}^{\boldsymbol{x}}.$$

It is worth mentioning that in most of the applications for $N_{n,k}$, $G_{n,k}$ and $M_{n,k}$, we have $\boldsymbol{l}_0 = \mathbf{0}$ and $\boldsymbol{\pi}_{\mathbf{0}} = (1, 0, \ldots, 0)$. In this case, we have $\boldsymbol{\varphi}_0(\boldsymbol{z}) = (1, 0, \ldots, 0)$.

Using Theorem 2.2, we can obtain the following marginal distribution of $\mathbf{X}_n$.

Suppose $\{i_1, \ldots, i_r\} \subset \{0, 1, \ldots, m\}$.

Let $\boldsymbol{X}_n^{(i_1,\ldots,i_r)} = (X_n^{(i_1)}, \ldots, X_n^{(i_r)})$, and let $\boldsymbol{x}^{(i_1,\ldots,i_r)} = (x^{(i_1)}, \ldots, x^{(i_r)})$ be a corresponding realization. Let $\boldsymbol{l}_t^{(i_1,\ldots,i_r)} = (l_t^{(i_1)}, \ldots, l_t^{(i_r)})$.

We introduce $\boldsymbol{g}_t(\boldsymbol{x}^{(i_1,\ldots,i_r)}) = \sum_{j\notin\{i_1,\ldots,i_r\}}\sum_{0\le x^{(j)}\le l_t^{(j)}} \boldsymbol{f}_t(\boldsymbol{x})$, $(t = 0, 1, \ldots, n)$.

Let $z_j = 1$, $j \neq i_1, \ldots, i_r$ in formula (2.7), we have the probability generating function of $\boldsymbol{X}_n^{(i_1,\ldots,i_r)}$

$$\varphi_n(z_1, \ldots, z_r) = \boldsymbol{\varphi}_0(z_1, \ldots, z_r)\prod_{t=1}^{n}\left[\left(A_t + \sum_{k\notin\{i_1,\ldots,i_r\}} B_t^{(k)}\right) + \sum_{j=1}^{r} z_{i_j}B_t^{(i_j)}\right]\mathbf{1}',$$

where

$$\boldsymbol{\varphi}_0(z_1, \ldots, z_r) = \sum_{0\le\boldsymbol{x}^{(i_1,\ldots,i_r)}\le\boldsymbol{l}_0^{(i_1,\ldots,i_r)}} \boldsymbol{g}_0(\boldsymbol{x}^{(i_1,\ldots,i_r)})z_1^{x_1}\cdots z_r^{x_r}.$$

And,

COROLLARY. *We have*

$$(2.8)\qquad \boldsymbol{g}_t(\boldsymbol{x}^{(i_1,\dots,i_r)}) = \boldsymbol{g}_{t-1}(\boldsymbol{x}^{(i_1,\dots,i_r)}) \left[A_t + \sum_{k \notin \{i_1,\dots,i_r\}} B_t^{(k)} \right] + \sum_{j=1}^{r} \boldsymbol{g}_{t-1}(\boldsymbol{x}^{(i_1,\dots,i_r)} - \boldsymbol{e}_j) B_t^{(i_j)} I(\boldsymbol{x}^{(i_1,\dots,i_r)} - \boldsymbol{e}_j \geq \boldsymbol{0}),$$

$$(\boldsymbol{0} \leq \boldsymbol{x}^{(i_1,\dots,i_r)} \leq \boldsymbol{l}_t^{(i_1,\dots,i_r)}), \quad (t = 0, 1, \dots, n),$$

and

$$(2.9)\qquad \Pr(\boldsymbol{X}_n^{(i_1,\dots,i_r)} = \boldsymbol{x}^{(i_1,\dots,i_r)}) = \boldsymbol{g}_n(\boldsymbol{x}^{(i_1,\dots,i_r)})\boldsymbol{1}', \quad \boldsymbol{0} \leq \boldsymbol{x}^{(i_1,\dots,i_r)} \leq \boldsymbol{l}_n^{(i_1,\dots,i_r)}.$$

The rest of the section will be used to the discussion of the homogeneous case, i.e. $A_t(\boldsymbol{x}) = A$, $B_t^{(k)}(\boldsymbol{x}) = B^{(k)}$, $(k = 0, 1, \dots, m)$, for all t and $\boldsymbol{x}$.

THEOREM 2.3. *If $A_t(\boldsymbol{x}) = A$, $B_t^{(k)}(\boldsymbol{x}) = B^{(k)}$, $(k = 0, 1, \dots, m)$, and $\boldsymbol{l}_0 = \boldsymbol{0}$, we have*

$$E(X_n^{(j)}) = \varphi_0(\boldsymbol{1}) \sum_{i=1}^{n} D^{i-1} B^{(j)} \boldsymbol{1}',$$

$$E(X_n^{(a)} X_n^{(b)}) = \varphi_0(\boldsymbol{1}) \sum_{i=1}^{n} \left[\sum_{s=1}^{i-1} D^{s-1} B^{(a)} D^{i-1-s} B^{(b)} + D^{i-1} B^{(b)} \sum_{s=1}^{n-i} D^{s-1} B^{(a)} \right] \boldsymbol{1}', \quad (a \neq b)$$

$$E((X_n^{(j)})^2) = \varphi_0(\boldsymbol{1}) \sum_{i=1}^{n} \left[\sum_{s=1}^{i-1} D^{s-1} B^{(j)} D^{i-1-s} + D^{i-1} B^{(j)} \sum_{s=1}^{n-i} D^{s-1} + D^{i-1} \right] B^{(j)} \boldsymbol{1}'.$$

And their generating functions are

$$M_{X^{(j)}}(w) = \sum_{n=1}^{\infty} E(X_n^{(j)}) w^n = \frac{w}{1-w} \varphi_0(\boldsymbol{1})(I - wD)^{-1} B^{(j)} \boldsymbol{1}',$$

$$M_{X^{(a)}X^{(b)}}(w) = \sum_{n=1}^{\infty} E(X_n^{(a)} X_n^{(b)}) w^n$$

$$= \frac{w^2}{1-w}\boldsymbol{\varphi}_0(\mathbf{1})(I-wD)^{-1}[B^{(a)}(I-wD)^{-1}B^{(b)} + B^{(b)}(I-wD)^{-1}B^{(a)}]\mathbf{1}', \qquad (a \neq b)$$

$$M_{(X^{(j)})^2}(w) = \sum_{n=1}^{\infty} E((X_n^{(j)})^2)w^n = \frac{w}{1-w}\boldsymbol{\varphi}_0(\mathbf{1})(I-wD)^{-1}B^{(j)}[2w(I-wD)^{-1}B^{(j)} + I]\mathbf{1}',$$

where $D = A + \sum_{k=0}^{m} B^{(k)}$, $(j, a, b = 0, 1, \ldots, m)$.

These formulas are complicated. But, noting the most run statistics satisfy the condition of the theorem, and we have that $\boldsymbol{\varphi}_0(\mathbf{1}) = (1, 0, \ldots, 0)$, and it is easy to apply these formulas.

PROOF. Taking the derivative in $\varphi(\boldsymbol{z})$ with respect to z_j and setting the $\boldsymbol{z} = \mathbf{1}$, and using $(A + \sum_{k=0}^{m} B^{(k)})\mathbf{1}' = \mathbf{1}'$ and

$$\frac{d}{dz}(A+zB)^n = \sum_{i=1}^{n}(A+zB)^{i-1}B(A+zB)^{n-i},$$

we can obtain these results.

Indeed, we have,

$$\begin{aligned}
E(X_n^{(j)}) &= \frac{\partial}{\partial z_j}\varphi_n(\boldsymbol{z})\mid_{\boldsymbol{z}=\mathbf{1}} = \frac{\partial}{\partial z_j}\left[\boldsymbol{\varphi}_0(\boldsymbol{z})\left(A + \sum_{k=0}^{m} z_k B^{(k)}\right)^n \mathbf{1}'\right]\Bigg|_{\boldsymbol{z}=\mathbf{1}} \\
&= \left(\frac{\partial}{\partial z_j}\boldsymbol{\varphi}_0(\boldsymbol{z})\mid_{\boldsymbol{z}=\mathbf{1}}\right)\left(A + \sum_{k=0}^{m} z_k B^{(k)}\right)^n \mathbf{1}'\mid_{\boldsymbol{z}=\mathbf{1}} \\
&\quad + \boldsymbol{\varphi}_0(\mathbf{1})\left(\frac{\partial}{\partial z_j}\left[A + \sum_{k=0}^{m} z_k B^{(k)}\right]^n\Bigg|_{\boldsymbol{z}=\mathbf{1}}\right)\mathbf{1}' \\
&= \left(\sum_{\mathbf{0}\le \boldsymbol{x}\le \boldsymbol{l}_0} x_j \boldsymbol{\pi}_{\boldsymbol{x}}\right)\left(A + \sum_{k=0}^{m} B^{(k)}\right)^n \mathbf{1}' \\
&\quad + \boldsymbol{\varphi}_0(\mathbf{1})\sum_{i=1}^{n}\left(A + \sum_{k=0}^{m} B^{(k)}\right)^{i-1} B^{(j)}\left(A + \sum_{k=0}^{m} B^{(k)}\right)^{n-i}\mathbf{1}' \\
&= \sum_{\mathbf{0}\le \boldsymbol{x}\le \boldsymbol{l}_0} x_j \boldsymbol{\pi}_{\boldsymbol{x}}\mathbf{1}' + \boldsymbol{\varphi}_0(\mathbf{1})\sum_{i=1}^{n} D^{i-1}B^{(j)}\mathbf{1}' = \boldsymbol{\varphi}_0(\mathbf{1})\sum_{i=1}^{n} D^{i-1}B^{(j)}\mathbf{1}',
\end{aligned}$$

and

$$\begin{aligned}
M_{X^{(j)}}(w) &= \sum_{n=1}^{\infty} E(X_n^{(j)})w^n = \boldsymbol{\varphi}_0(\mathbf{1})\left[\sum_{n=1}^{\infty}\sum_{i=1}^{n} D^{i-1}w^n\right] B^{(j)}\mathbf{1}' \\
&= \boldsymbol{\varphi}_0(\mathbf{1})\left[\sum_{i=1}^{\infty}\sum_{n=i}^{\infty} D^{i-1}w^n\right] B^{(j)}\mathbf{1}'
\end{aligned}$$

$$= \varphi_0(\mathbf{1}) \left[\sum_{i=1}^{\infty} \left[D^{i-1} \left(\sum_{n=i}^{\infty} w^n \right) \right] \right] B^{(j)} \mathbf{1}'$$

$$= \frac{w}{1-w} \varphi_0(\mathbf{1}) \left[\sum_{i=1}^{\infty} (wD)^{i-1} \right] B^{(j)} \mathbf{1}' = \frac{w}{1-w} \varphi_0(\mathbf{1}) [I - wD]^{-1} B^{(j)} \mathbf{1}'.$$

When $a \neq b$ $(0 \leq a, b \leq m)$, we have

$$E(X_n^{(a)} X_n^{(b)})$$

$$= \frac{\partial^2}{\partial z_a \partial z_b} \varphi_n(\mathbf{z}) \mid_{\mathbf{z}=\mathbf{1}} = \left(\frac{\partial^2}{\partial z_a \partial z_b} \varphi_0(\mathbf{z}) \mid_{\mathbf{z}=\mathbf{1}} \right) \left(A + \sum_{k=0}^{m} B^{(k)} \right)^n \mathbf{1}'$$

$$+ \left(\frac{\partial}{\partial z_b} \varphi_0(\mathbf{z}) \mid_{\mathbf{z}=\mathbf{1}} \right) \sum_{i=1}^{n} D^{i-1} B^{(a)} \mathbf{1}' + \left(\frac{\partial}{\partial z_a} \varphi_0(\mathbf{z}) \mid_{\mathbf{z}=\mathbf{1}} \right) \sum_{i=1}^{n} D^{i-1} B^{(b)} \mathbf{1}'$$

$$+ \varphi_0(\mathbf{1}) \sum_{i=1}^{n} \left[\sum_{s=1}^{i-1} D^{s-1} B^{(a)} D^{i-1-s} B^{(b)} + D^{i-1} B^{(b)} \sum_{s=1}^{n-i} D^{s-1} B^{(a)} \right] \mathbf{1}'$$

$$= \varphi_0(\mathbf{1}) \sum_{i=1}^{n} \left[\sum_{s=1}^{i-1} D^{s-1} B^{(a)} D^{i-1-s} B^{(b)} + D^{i-1} B^{(b)} \sum_{s=1}^{n-i} D^{s-1} B^{(a)} \right] \mathbf{1}',$$

and

$$M_{X^{(a)} X^{(b)}}(w) = \sum_{n=1}^{\infty} E(X_n^{(a)} X_n^{(b)}) w^n$$

$$= \varphi_0(\mathbf{1}) \left[\sum_{n=1}^{\infty} w^n \left[\sum_{i=1}^{n} \sum_{s=1}^{i-1} D^{s-1} B^{(a)} D^{i-1-s} B^{(b)} \right.\right.$$

$$\left.\left. + \sum_{i=1}^{n} \sum_{s=1}^{n-i} D^{i-1} B^{(b)} D^{s-1} B^{(a)} \right] \right] \mathbf{1}',$$

where

$$\sum_{n=1}^{\infty} \sum_{i=1}^{n} \sum_{s=1}^{i-1} D^{s-1} B^{(a)} D^{i-1-s} w^n$$

$$= \sum_{n=1}^{\infty} \sum_{s=1}^{n} \sum_{i=s+1}^{n} D^{s-1} B^{(a)} D^{i-1-s} w^n$$

$$= \sum_{n=1}^{\infty} \sum_{s=1}^{n} \sum_{(i-s)=1}^{n-s} D^{s-1} B^{(a)} D^{(i-s)-1} w^n = \sum_{n=1}^{\infty} \sum_{s=1}^{n} \sum_{r=1}^{n-s} D^{s-1} B^{(a)} D^{r-1} w^n$$

$$= \sum_{n=1}^{\infty} \sum_{s=1}^{n} D^{s-1} B^{(a)} \left(\sum_{r=1}^{n-s} D^{r-1} w^n \right) = \sum_{s=1}^{\infty} \sum_{n=s}^{\infty} D^{s-1} B^{(a)} \left(\sum_{r=1}^{n-s} D^{r-1} w^n \right)$$

$$= \sum_{s=1}^{\infty} D^{s-1} B^{(a)} \left(\sum_{n=s}^{\infty} \sum_{r=1}^{n-s} D^{r-1} w^n \right) = \sum_{s=1}^{\infty} D^{s-1} B^{(a)} \left(\sum_{r=1}^{\infty} \sum_{n=r+s}^{\infty} D^{r-1} w^n \right)$$

$$= \sum_{s=1}^{\infty} D^{s-1}B^{(a)}\left[\sum_{r=1}^{\infty} D^{r-1}\left(\sum_{n=r+s}^{\infty} w^n\right)\right] = \sum_{s=1}^{\infty} D^{s-1}B^{(a)}\left[\sum_{r=1}^{\infty} D^{r-1}\frac{w^{r+s}}{1-w}\right]$$

$$= \frac{w^2}{1-w}\sum_{s=1}^{\infty}\left(D^{s-1}B^{(a)}\left[\sum_{r=1}^{\infty} D^{r-1}w^{r-1}\right]w^{s-1}\right)$$

$$= \frac{w^2}{1-w}\left(\sum_{s=1}^{\infty} D^{s-1}w^{s-1}\right)B^{(a)}\left(\sum_{r=1}^{\infty} D^{r-1}w^{r-1}\right)$$

$$= \frac{w^2}{1-w}(I-wD)^{-1}B^{(a)}(I-wD)^{-1}.$$

Similarly, we have

$$\sum_{n=1}^{\infty}\sum_{i=1}^{n}\sum_{s=1}^{n-i} D^{i-1}B^{(b)}D^{s-1}w^n = \frac{w^2}{1-w}(I-wD)^{-1}B^{(b)}(I-wD)^{-1}.$$

Hence we get

$$M_{X^{(a)}X^{(b)}}(w)$$
$$= \frac{w^2}{1-w}\varphi_0(\mathbf{1})(I-wD)^{-1}[B^{(a)}(I-wD)^{-1}B^{(b)} + B^{(b)}(I-wD)^{-1}B^{(a)}]\mathbf{1}'.$$

Similarly, we can deduce the $E((X_n^{(j)})^2)$ and $M_{(X^{(j)})^2}(w)$.
This completes the proof.

We introduce the double vector generating function $\Phi(z_0, z_1, \ldots, z_m, w) = \sum_{t=0}^{\infty}\varphi_t(z_0, \ldots, z_m)w^t$, for homogeneous case, then $\Phi(\boldsymbol{z}, w) = \varphi_0(\boldsymbol{z})[I - w(A + \sum_{k=0}^{m} z_k B^{(k)})]^{-1}$. (Because $\Phi(\boldsymbol{z}, w) = \sum_{t=0}^{\infty}[\varphi_0(\boldsymbol{z})(A + \sum_{k=0}^{m} z_k B^{(k)})^t w^t] = \varphi_0(\boldsymbol{z})[\sum_{t=0}^{\infty}(w(A + \sum_{k=0}^{m} z_k B^{(k)}))^t] = \varphi_0(\boldsymbol{z})[I - w(A + \sum_{k=0}^{m} z_k B^{(k)})]^{-1}$.)

3. Joint distributions of $N_{n,k_i}^{(i)}$, $M_{n,k_i}^{(i)}$ and $G_{n,k_i}^{(i)}$ runs

Suppose we are given a sequence of trials $Y_1, Y_2, \ldots$, with possible outcomes "0", "1", …, "m", and $\Pr(Y_t = i) = p_t^{(i)}$, $(i = 0, 1, \ldots, m)$, $t \geq 1$.

Let us denote by x_i the number of "i"-runs of length k_i up to the t-th trial $(i = 0, 1, \ldots, m)$. Let y_i^* be the number of trailing "i", i.e. the number of last consecutive "i" counting backwards. Clearly, only one of $y_0^*, y_1^*, \ldots, y_m^*$ is not equal to 0.

3.1 *Non-overlapping runs* $(N_{n,k_0}^{(0)}, N_{n,k_1}^{(1)}, \ldots, N_{n,k_m}^{(m)})$

Let $l_n^{(i)} = [\frac{n}{k_i}]$, and $y_i = y_i^* - [\frac{y_i^*}{k_i}]k_i$, $(i = 0, 1, \ldots, m)$.
We can give a Markov chain $\{Z_t = (x_0, x_1, \ldots, x_m; y_0, y_1, \ldots, y_m) \mid t \geq 0\}$.
The state space is

$$\Omega = \{(x_0, x_1, \ldots, x_m; y_0, y_1, \ldots, y_m) \mid 0 \leq x_i \leq l_n^{(i)}, 0 \leq y_i \leq k_i - 1,$$
$$i = 0, 1, \ldots, m\}.$$

Define $\boldsymbol{U_x} = \{(\boldsymbol{x}; y_0, y_1, \dots, y_m) \mid 0 \le y_i \le k_i - 1, i = 0, 1, \dots, m\}$, then $\Omega = \bigcup_{0 \le \boldsymbol{x} \le \boldsymbol{l}_n} \boldsymbol{U_x}$.

Using lexicographical order, we let $\boldsymbol{U_x} = \{(\boldsymbol{x}; 0, \dots, 0), (\boldsymbol{x}; 0, \dots, 1), \dots, (\boldsymbol{x}; 0, \dots, k_m - 1), (\boldsymbol{x}; 0, \dots, 1, 0), (\boldsymbol{x}; 0, \dots, 2, 0), \dots, (\boldsymbol{x}; 0, \dots, k_{m-1} - 1, 0), \dots, (\boldsymbol{x}; 1, 0, \dots, 0), \dots, (\boldsymbol{x}; k_0 - 1, 0, \dots, 0)\} \equiv \{U_{\boldsymbol{x},1}, U_{\boldsymbol{x},2}, \dots, U_{\boldsymbol{x},s}\}$, where $s = \sum_{i=0}^{m}(k_i - 1) + 1$.

With this set up, the random vector $(N_{n,k_0}^{(0)}, N_{n,k_1}^{(1)}, \dots, N_{n,k_m}^{(m)})$ satisfies Definition 2.1, with

$$\Pr(Z_t = (\boldsymbol{x}; 0, \dots, y_i + 1, 0, \dots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \dots, y_i, 0, \dots, 0)) = p_t^{(i)} \quad 0 \le y_i \le k_i - 2, \quad 0 \le i \le m, \tag{3.1}$$

$$\Pr(Z_t = (\boldsymbol{x}; 0, \dots, \overbrace{1}^{(j)}, \dots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \dots, y_i, \dots, 0)) = p_t^{(j)} \quad 0 \le y_i \le k_i - 1, \quad 0 \le i, j \le m, \quad i \ne j, \tag{3.2}$$

$$\Pr(Z_t = (\boldsymbol{x} + \boldsymbol{e}_i; 0, \dots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \dots, k_i - 1, 0, \dots, 0)) = p_t^{(i)} \quad 0 \le i \le m. \tag{3.3}$$

Matrices $A_t(\boldsymbol{x})$, $B_t^{(k)}(\boldsymbol{x})$ $(k = 0, 1, \dots, m)$ can be given by formulas (3.1), (3.2) and (3.3) as formulas (2.1) and (2.2).

As an illustration, let $m = 2$, $k_0 = 2$, $k_1 = 2$, $k_2 = 3$, then we have $s = (2 + 2 + 3) - 2 = 5$, and $\boldsymbol{U_x} = \{(\boldsymbol{x}; 0, 0, 0), (\boldsymbol{x}; 0, 0, 1), (\boldsymbol{x}; 0, 0, 2), (\boldsymbol{x}; 0, 1, 0), (\boldsymbol{x}; 1, 0, 0)\}$, with

$$A_t(\boldsymbol{x}) = \begin{pmatrix} (\boldsymbol{x}; 0,0,0) & (\boldsymbol{x}; 0,0,1) & (\boldsymbol{x}; 0,0,2) & (\boldsymbol{x}; 0,1,0) & (\boldsymbol{x}; 1,0,0) \\ 0 & p_t^{(2)} & 0 & p_t^{(1)} & p_t^{(0)} \\ 0 & 0 & p_t^{(2)} & p_t^{(1)} & p_t^{(0)} \\ 0 & 0 & 0 & p_t^{(1)} & p_t^{(0)} \\ 0 & p_t^{(2)} & 0 & 0 & p_t^{(0)} \\ 0 & p_t^{(2)} & 0 & p_t^{(1)} & 0 \end{pmatrix},$$

$$B_t^{(2)}(\boldsymbol{x}) = \begin{pmatrix} (\boldsymbol{x} + \boldsymbol{e}_2; 0,0,0) & (\cdot; 0,0,1) & (\cdot; 0,0,2) & (\cdot; 0,1,0) & (\cdot; 1,0,0) \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ p_t^{(2)} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$B_t^{(1)}(\boldsymbol{x}) = \begin{pmatrix} (\boldsymbol{x} + \boldsymbol{e}_1; 0,0,0) & (\cdot; 0,0,1) & (\cdot; 0,0,2) & (\cdot; 0,1,0) & (\cdot; 1,0,0) \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ p_t^{(1)} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$B_t^{(0)}(\boldsymbol{x}) = \begin{pmatrix} (\boldsymbol{x}+\boldsymbol{e}_0;0,0,0) & (\cdot;0,0,1) & (\cdot;0,0,2) & (\cdot;0,1,0) & (\cdot;1,0,0) \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ p_t^{(0)} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Therefore the probability function of the random vector $(N_{t,k_0}^{(0)}, N_{t,k_1}^{(1)}, \dots, N_{t,k_m}^{(m)})$ can be successively evaluated for all t by making use of Theorem 2.1.

Clearly, $A_t(\boldsymbol{x}) = A_t$, $B_t^{(k)}(\boldsymbol{x}) = B_t^{(k)}$, $(k = 0, 1, \dots, m)$, so Theorem 2.2 can be used. Further, in the i.i.d. case or the Markov chain dependent trial case, we have $A_t(\boldsymbol{x}) = A$, $B_t^{(k)} = B^{(k)}$ $(k = 0, \dots, m)$, Theorem 2.3 can be used.

3.2 *Overlapping runs* $(M_{n,k_0}^{(0)}, M_{n,k_1}^{(1)}, \dots, M_{n,k_m}^{(m)})$

Let $l_n^{(i)} = n - k_i + 1$, and

$$y_i = \begin{cases} y_i^* & \text{if} \quad y_i^* \le k_i - 1 \\ -1 & \text{if} \quad y_i^* \ge k_i \end{cases}.$$

Then, we have

$$\begin{aligned} \Omega &= \{(\boldsymbol{x}; y_0, y_1, \dots, y_m) \mid \boldsymbol{0} \le \boldsymbol{x} \le \boldsymbol{l}_n, -1 \le y_i \le k_i - 1, i = 0, 1, \dots, m\} \\ &= \bigcup_{\boldsymbol{0} \le \boldsymbol{x} \le \boldsymbol{l}_n} \boldsymbol{U}_{\boldsymbol{x}}, \end{aligned}$$

where $\boldsymbol{U}_{\boldsymbol{x}} = \{(\boldsymbol{x}; y_0, y_1, \dots, y_m) \mid -1 \le y_i \le k_i - 1, i = 0, 1, \dots, m\} = \{(\boldsymbol{x}; 0, \dots, 0), (\boldsymbol{x}; 0, \dots, 1), \dots, (\boldsymbol{x}; 0, \dots, k_m - 1), (\boldsymbol{x}; 0, \dots, -1), (\boldsymbol{x}; 0, \dots, 1, 0), (\boldsymbol{x}; 0, \dots, 2, 0), \dots, (\boldsymbol{x}; 0, \dots, k_{m-1} - 1, 0), (\boldsymbol{x}, 0, \dots, -1, 0), \dots, (\boldsymbol{x}; 1, 0, \dots, 0), \dots, (\boldsymbol{x}; k_0 - 1, 0, \dots, 0), (\boldsymbol{x}, -1, 0, \dots, 0)\} \equiv \{U_{\boldsymbol{x},1}, U_{\boldsymbol{x},2}, \dots, U_{\boldsymbol{x},s}\}$, and $s = \sum_{i=0}^m k_i + 1$. We have

$$\Pr(Z_t = (\boldsymbol{x}; 0, \dots, y_i + 1, 0, \dots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \dots, y_i, 0, \dots, 0)) = p_t^{(i)} \qquad 0 \le y_i \le k_i - 2, \quad 0 \le i \le m,$$

$$\Pr(Z_t = (\boldsymbol{x}; 0, \dots, \overbrace{1}^{(j)}, \dots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \dots, y_i, \dots, 0)) = p_t^{(j)} \qquad -1 \le y_i \le k_i - 1, \quad 0 \le i, j \le m, \quad i \ne j,$$

$$\Pr(Z_t = (\boldsymbol{x} + \boldsymbol{e}_i; 0, \dots, \overbrace{-1}^{(i)}, \dots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \dots, k_i - 1, 0, \dots, 0)) = p_t^{(i)} \qquad 0 \le i \le m, i,$$

$$\Pr(Z_t = (\boldsymbol{x} + \boldsymbol{e}_i; 0, \dots, \overbrace{-1}^{(i)}, \dots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \dots, \overbrace{-1}^{(i)}, \dots, 0)) = p_t^{(i)} \qquad 0 \le i \le m.$$

For the above illustration, we have $s = 8$, and $\boldsymbol{U}_{\boldsymbol{x}} = \{(\boldsymbol{x};0,0,0),(\boldsymbol{x};0,0,1),(\boldsymbol{x};0,0,2),(\boldsymbol{x};0,0,-1),(\boldsymbol{x};0,1,0),(\boldsymbol{x},0,-1,0),(\boldsymbol{x};1,0,0),(\boldsymbol{x};-1,0,0)\}$, with

$$A_t(\boldsymbol{x}) = \begin{pmatrix} (000) & (001) & (002) & (00-1) & (010) & (0-10) & (100) & (-100) \\ 0 & p_t^{(2)} & 0 & 0 & p_t^{(1)} & 0 & p_t^{(0)} & 0 \\ 0 & 0 & p_t^{(2)} & 0 & p_t^{(1)} & 0 & p_t^{(0)} & 0 \\ 0 & 0 & 0 & 0 & t_t^{(1)} & 0 & p_t^{(0)} & 0 \\ 0 & 0 & 0 & 0 & p_t^{(1)} & 0 & p_t^{(0)} & 0 \\ 0 & p_t^{(2)} & 0 & 0 & 0 & 0 & p_t^{(0)} & 0 \\ 0 & p_t^{(2)} & 0 & 0 & 0 & 0 & p_t^{(0)} & 0 \\ 0 & p_t^{(2)} & 0 & 0 & p_t^{(1)} & 0 & 0 & 0 \\ 0 & p_t^{(2)} & 0 & 0 & p_t^{(1)} & 0 & 0 & 0 \end{pmatrix},$$

$$B_t^{(2)}(\boldsymbol{x}) = \begin{pmatrix} (000) & (001) & (002) & (00-1) & (010) & (0-10) & (100) & (-100) \\ & & & 0 & & & & \\ & & & 0 & & & & \\ & & & p_t^{(2)} & & & & \\ & & & p_t^{(2)} & & & & \\ & & & 0 & & & & \\ & & & 0 & & & & \\ & & & 0 & & & & \\ & & & 0 & & & & \end{pmatrix},$$

$$B_t^{(1)}(\boldsymbol{x}) = \begin{pmatrix} (000) & (001) & (002) & (00-1) & (010) & (0-10) & (100) & (-100) \\ & & & & & 0 & & \\ & & & & & 0 & & \\ & & & & & 0 & & \\ & & & & & 0 & & \\ & & & & & p_t^{(1)} & & \\ & & & & & p_t^{(1)} & & \\ & & & & & 0 & & \\ & & & & & 0 & & \end{pmatrix},$$

$$B_t^{(0)}(\boldsymbol{x}) = \begin{pmatrix} (000) & (001) & (002) & (00-1) & (010) & (0-10) & (100) & (-100) \\ & & & & & & & 0 \\ & & & & & & & 0 \\ & & & & & & & 0 \\ & & & & & & & 0 \\ & & & & & & & 0 \\ & & & & & & & 0 \\ & & & & & & & p_t^{(0)} \\ & & & & & & & p_t^{(0)} \end{pmatrix}.$$

3.3 *Runs of length at least* k_i, $(G^{(0)}_{n,k_0}, G^{(1)}_{n,k_1}, \ldots, G^{(m)}_{n,k_m})$

Let $l^{(i)}_n = [\frac{n+1}{k_i+1}]$, and

$$y_i = \begin{cases} y_i^* & \text{if } \ y_i^* \le k_i - 1 \\ -1 & \text{if } \ y_i^* \ge k_i \end{cases} .$$

The state space is

$$\begin{aligned}\Omega &= \{(\boldsymbol{x}; y_0, y_1, \ldots, y_m) \mid \boldsymbol{0} \le \boldsymbol{x} \le \boldsymbol{l}_n, -1 \le y_i \le k_i - 1, i = 0, 1, \ldots, m\} \\ &= \bigcup_{\boldsymbol{0} \le \boldsymbol{x} \le \boldsymbol{l}_n} \boldsymbol{U}_{\boldsymbol{x}},\end{aligned}$$

where $\boldsymbol{U}_{\boldsymbol{x}} = \{(\boldsymbol{x}; y_0, y_1, \ldots, y_m) \mid -1 \le y_i \le k_i - 1, i = 0, 1, \ldots, m\}$, and $s = \sum_{i=0}^{m} k_i + 1$. And we have

$$\Pr(Z_t = (\boldsymbol{x}; 0, \ldots, y_i + 1, 0, \ldots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \ldots, y_i, 0, \ldots, 0)) = p_t^{(i)} \qquad 0 \le y_i \le k_i - 2, \quad 0 \le i \le m,$$

$$\Pr(Z_t = (\boldsymbol{x}; 0, \ldots, \overbrace{1}^{(j)}, \ldots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \ldots, y_i, \ldots, 0)) = p_t^{(j)} \qquad -1 \le y_i \le k_i - 1, \quad 0 \le i, j \le m, \quad i \ne j,$$

$$\Pr(Z_t = (\boldsymbol{x} + \boldsymbol{e}_i; 0, \ldots, \overbrace{-1}^{(i)}, \ldots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \ldots, k_i - 1, 0, \ldots, 0)) = p_t^{(i)} \qquad 0 \le i \le m,$$

and

$$\Pr(Z_t = (\boldsymbol{x}; 0, \ldots, \overbrace{-1}^{(i)}, \ldots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \ldots, \overbrace{-1}^{(i)}, \ldots, 0)) = p_t^{(i)} \qquad 0 \le i \le m.$$

3.4 *Numerical example*

We consider the distribution of $(N^{(0)}_{5,2}, N^{(1)}_{5,2}, N^{(2)}_{5,3})$ in a sequence of three-state trials of length $n = 5$. For outcomes "0", "1", "2", we assume that the t-th trial's probability is $p_t^{(0)} = \frac{1}{t+1}$, $p_t^{(1)} = \frac{1}{t+2}$, $p_t^{(2)} = 1 - p_t^{(0)} - p_t^{(1)}$.

Using Theorem 2.2, where A_t and $B_t^{(k)}$ $(k = 0, 1, 2)$ are given in Subsection 3.1, we get $\varphi_5(z_0, z_1, z_2) = 0.01389z_0^2 + 0.01708z_0z_1 + 0.04136z_0z_2 + 0.19017z_0 + 0.00516z_1^2 + 0.02095z_1z_2 + 0.11593z_1 + 0.23713z_2 + 0.35832$. Hence Table 1 gives the exact joint distribution of $(N^{(0)}_{5,2}, N^{(1)}_{5,2}, N^{(2)}_{5,3})$.

Let $z_1 = 1$ and $z_2 = 1$, we have $\varphi_5(z_0, z_1 = 1, z_2 = 1) = 0.01389z_0^2 + 0.24864z_0 + 0.73749$. The marginal distribution of $N^{(0)}_{5,2}$ is $\Pr(N^{(0)}_{5,2} = 2) = 0.01389$, $\Pr(N^{(0)}_{5,2} = 1) = 0.24864$, and $\Pr(N^{(0)}_{5,2} = 0) = 0.73749$. This is the example of Fu and Koutras (1994).

Table 1. The exact joint distribution of $(N_{5,2}^{(0)}, N_{5,2}^{(1)}, N_{5,2}^{(2)})$.

		$X_1 = 0$	$X_1 = 1$	$X_1 = 2$	
$X_0 = 0$	$X_2 = 0$	0.35832	0.11593	0.00516	
	$X_2 = 1$	0.23713	0.02095		0.73749
$X_0 = 1$	$X_2 = 0$	0.19017	0.01708		
	$X_2 = 1$	0.04136			0.24864
$X_0 = 2$	$X_2 = 0$	0.01389			
	$X_2 = 1$				0.01389

4. Markov chain imbeddable variable of returnable type

In this section, we prove that Markov chains which can move backwards to the neighboring state possess the similar property.

Let X_n be a nonnegative integer valued random variable, and $l_n = \max\{x : \Pr(X_n = x) > 0\}$.

DEFINITION 4.1. The nonnegative integer valued random variable X_n is called a Markov chain imbeddable variable of returnable type(M.V.R.), if

(1) there exists a Markov chain $\{Z_t, t \geq 0\}$ defined on a finite state space Ω,
(2) there exists a partition $\{U_x, x = 0, 1, \dots\}$ on the state space Ω,
(3) for every $x = 0, 1, \dots$, we have

$$\Pr(X_n = x) = \Pr(Z_n \in U_x), \qquad x = 0, 1, \dots$$

and (4)

$$\Pr(Z_t \in U_y \mid Z_{t-1} \in U_x) = 0, \qquad \text{if} \quad y \neq x - 1, x, x + 1.$$

Without loss of generality, we assume U_x $(x = 0, 1, \dots)$ have the same cardinality $s = |U_x|$, and denote $U_x = \{u_{x1}, u_{x2}, \dots, u_{xs}\}$.

We introduce the probability vectors

$$\boldsymbol{f}_t(x) = (\Pr(Z_t = u_{x1}), \Pr(Z_t = u_{x2}), \dots, Pr(Z_t = u_{xs})),$$
$$x = 0, 1, \dots, l_t, \qquad t \geq 0,$$

where l_t is the largest subscript of partitions U_x which include accessible state of the t-th step Z_t of the Markov chain, i.e.

$$l_t = \max\{x : \Pr(Z_t \in U_x) > 0\},$$

when $t = n$, it is l_n. Let $\boldsymbol{\pi}_x$ be the initial probabilities of the Markov chain $\{Z_t, t \geq 0\}$, i.e. $\boldsymbol{\pi}_x = \boldsymbol{f}_0(x)$, $(x = 0, 1, \dots, l_t)$.

We introduce $s \times s$ transition probability matrices

$$A_t(x) = (\Pr(Z_t = u_{xj} \mid Z_{t-1} = u_{xi}))_{s\times s},$$
$$B_t(x) = (\Pr(Z_t = u_{x+1,j} \mid Z_{t-1} = u_{xi}))_{s\times s},$$
$$C_t(x) = (\Pr(Z_t = u_{x-1,j} \mid Z_{t-1} = u_{xi}))_{s\times s}.$$

Clearly, (1) the entries of $A_t(x)$ control the within state one-step transitions, (2) the entries of $B_t(x)$ control the one-step transitions from state U_x to state U_{x+1}, (3) $C_t(x)$ control the return transitions from U_x to U_{x-1}.

Using the Chapman-Kolmogorov equations, we have

THEOREM 4.1. *The probability vectors* $\boldsymbol{f}_t(x)$, $0 \le x \le l_t$, $t = 0, 1, \ldots, n$, *satisfy*

$$\boldsymbol{f}_t(x) = \boldsymbol{f}_{t-1}(x)A_t(x) + \boldsymbol{f}_{t-1}(x-1)B_t(x-1) + \boldsymbol{f}_{t-1}(x+1)xC_t(x+1) \quad (x = 0, 1, \ldots, l_t), \quad (t = 1, 2, \ldots, n), \tag{4.1}$$

and

$$\boldsymbol{f}_t(y) = \mathbf{0}, \quad y < 0 \quad or \quad y > l_t, \quad (t = 1, 2, \ldots, n),$$

and $\boldsymbol{f}_0(x) = \boldsymbol{\pi}_x$, $(x = 0, 1, \ldots, l_0)$.

The probability distribution function of a M.V.R. X_n *is given by*

$$\Pr(X_n = x) = \boldsymbol{f}_n(x)\mathbf{1}', \quad (x = 0, 1, \ldots, l_n). \tag{4.2}$$

We consider the probability generating function of X_n,

$$\varphi_n(z) = \sum_{x=0}^{l_n} \Pr(X_n = x)z^x = \sum_{x=0}^{l_n} \boldsymbol{f}_n(x)\mathbf{1}'z^x = \left(\sum_{x=0}^{l_n} \boldsymbol{f}_n(x)z^x\right)\mathbf{1}'.$$

We define $\boldsymbol{\varphi}_t(z) = \sum_{x=0}^{l_t} \boldsymbol{f}_t(x)z^x$ for the Markov chain, then $\varphi_n(z) = \boldsymbol{\varphi}_n(z)\mathbf{1}'$.

THEOREM 4.2. *If* $A_t(x) = A_t$, $B_t(x) = B_t$, $C_t(x) = C_t$ *for all* $x = 0, 1, \ldots,$ *(i.e. the transition probabilities do not depend on* x*), we have*

$$\boldsymbol{\varphi}_t(z) = \boldsymbol{\varphi}_{t-1}(z)(A_t + zB_t + z^{-1}C_t), \quad t = 1, 2, \ldots, \tag{4.3}$$

where $\boldsymbol{\varphi}_0(z) = \sum_{x=0}^{l_0} \boldsymbol{\pi}_x z^x$.

PROOF.

$$\begin{aligned}\boldsymbol{\varphi}_t(z) &= \sum_{x=0}^{l_t} \boldsymbol{f}_t(x)z^x \\ &= \sum_{x=0}^{l_t} [\boldsymbol{f}_{t-1}(x)A_t + \boldsymbol{f}_{t-1}(x-1)B_t + \boldsymbol{f}_{t-1}(x+1)C_t]z^x \\ &= \sum_{x=0}^{l_t} \boldsymbol{f}_{t-1}(x)A_t z^x + \sum_{x=0}^{l_t-1} \boldsymbol{f}_{t-1}(x)B_t z^{x+1} + \sum_{x=1}^{l_t+1} \boldsymbol{f}_{t-1}(x)C_t z^{x-1}.\end{aligned}$$

Because $l_t = \max\{x : \Pr(Z_t \in U_x) > 0\}$, we have $\boldsymbol{f}_t(x) = \mathbf{0}$, $x > l_t$ and $0 \le l_0 \le l_1 \le \cdots \le l_n$. We get $l_t - l_{t-1} \in \{0, 1\}$, since the Markov chain cannot jump directly to a higher state without visiting its next state.

(1) If $l_t = l_{t-1}$, we have

$$\begin{aligned}\boldsymbol{\varphi}_t(z) = &\left(\sum_{x=0}^{l_{t-1}} \boldsymbol{f}_{t-1}(x)z^x\right) A_t + \left(\sum_{x=0}^{l_{t-1}} \boldsymbol{f}_{t-1}(x)z^x - \boldsymbol{f}_{t-1}(l_{t-1})z^{l_{t-1}}\right) zB_t \\ &+ \left(\sum_{x=0}^{l_{t-1}} \boldsymbol{f}_{t-1}(x)z^x - \boldsymbol{f}_{t-1}(0)z^0 + \boldsymbol{f}_{t-1}(l_{t-1}+1)C_t z^{l_{t-1}}\right) z^{-1}C_t \\ &= \boldsymbol{\varphi}_{t-1}(z)(A_t + zB_t + z^{-1}C_t) - \boldsymbol{f}_{t-1}(l_{t-1})z^{l_{t-1}+1}B_t - \boldsymbol{f}_{t-1}(0)z^{-1}C_t + \mathbf{0}.\end{aligned}$$

Setting $z = 1$ and multiplying both sides by $\mathbf{1}'$, because of $\boldsymbol{\varphi}_t(1)\mathbf{1}' = \varphi_t(1) = \sum_{x=0}^{l_t} \Pr(X_t = x) = 1$, and $(A_t + B_t + C_t)\mathbf{1}' = \mathbf{1}'$, $\boldsymbol{\varphi}_{t-1}(1)\mathbf{1}' = 1$, we obtain

$$(\boldsymbol{f}_{t-1}(l_{t-1})B_t + \boldsymbol{f}_{t-1}(0)C_t)\mathbf{1}' = 0.$$

Because every term in the left hand side of the above formula is nonnegative, we get

$$\boldsymbol{f}_{t-1}(l_{t-1})B_t = \mathbf{0} \quad \text{and} \quad \boldsymbol{f}_{t-1}(0)C_t = \mathbf{0}.$$

Hence we have

$$\boldsymbol{\varphi}_t(z) = \boldsymbol{\varphi}_{t-1}(z)(A_t + zB_t + z^{-1}C_t).$$

(2) If $l_t = l_{t-1} + 1$, we have

$$\begin{aligned}\boldsymbol{\varphi}_t(z) = &\left(\sum_{x=0}^{l_{t-1}} \boldsymbol{f}_{t-1}(x)z^x + \boldsymbol{f}_{t-1}(l_{t-1}+1)z^{l_{t-1}+1}\right) A_t + \left(\sum_{x=0}^{l_{t-1}} \boldsymbol{f}_{t-1}(x)z^x\right) zB_t \\ &+ \left(\sum_{x=0}^{l_{t-1}} \boldsymbol{f}_{t-1}(x)z^x - \boldsymbol{f}_{t-1}(0)z^0 + \sum_{i=1}^{2} \boldsymbol{f}_{t-1}(l_{t-1}+i)z^{l_{t-1}+i}\right) z^{-1}C_t \\ &= \boldsymbol{\varphi}_{t-1}(A_t + zB_t + z_{-1}C_t) + \boldsymbol{f}_{t-1}(l_{t-1}+1)z^{l_{t-1}+1}A_t \\ &\quad - \boldsymbol{f}_{t-1}(0)z^{-1}C_t + \boldsymbol{f}_{t-1}(l_{t-1}+1)z^{l_{t-1}}C_t + \boldsymbol{f}_{t-1}(l_{t-1}+2)z^{l_{t-1}+1}C_t \\ &= \boldsymbol{\varphi}_{t-1}(z)(A_t + zB_t + z^{-1}C_t) + \mathbf{0} - \boldsymbol{f}_{t-1}(0)z^{-1}C_t + \mathbf{0} + \mathbf{0},\end{aligned}$$

where $\boldsymbol{f}_{t-1}(l_{t-1}+1) = \mathbf{0}$, $\boldsymbol{f}_{t-1}(l_{t-1}+2) = \mathbf{0}$ and the term $\boldsymbol{f}_{t-1}(0)C_t = \mathbf{0}$ by the same argument as (1).

Therefore, we get

$$\boldsymbol{\varphi}_t(z) = \boldsymbol{\varphi}_{t-1}(z)(A_t + zB_t + z^{-1}C_t), \qquad t \geq 0.$$

This completes the proof.

Remark. The $\boldsymbol{f}_{t-1}(x)C_t$ is the backwards transition probability vector from U_x to U_{x-1}, and the $\boldsymbol{f}_{t-1}(x)B_t$ is the transition probability vector from U_x to U_{x+1}. Because the random variable X_t is nonnegative and $l_t = \max\{x : \Pr(Z_t \in U_x) > 0\}$ is its upper end point, these transition can not occur when $x = 0$ or $x = l_{t-1}$ (at

$l_t = l_{t-1}$). This is an intrinsic reason for $\boldsymbol{f}_{t-1}(0)C_t = \mathbf{0}$ and $\boldsymbol{f}_{t-1}(l_{t-1})B_t = \mathbf{0}$ (at $l_t = l_{t-1}$). By the reason, Theorem 2.3 can be viewed intuitively.

In fact, Theorem 4.2 can be recognized by the following reason. The coefficient vector $\boldsymbol{f}_{t-1}(x)$ of z^x of $\boldsymbol{\varphi}_{t-1}(z)$ reflects the probability vector of $\{Z_{t-1} \in U_x\}$. Given $\{Z_{t-1} \in U_x\}$, there are three possible transitions: to U_x, U_{x+1} and U_{x-1}. Given $\{Z_{t-1} \in U_x\}$, the probability vector of $\{Z_t \in U_x\}$ is $\boldsymbol{f}_{t-1}(x)A_t$, which is a part of coefficient of z^x of $\boldsymbol{\varphi}_t(z)$; the probability vector of $\{Z_t \in U_{x+1}\}$ is $\boldsymbol{f}_{t-1}(x)B_t$, which is a part of coefficient of $z^{(x+1)}$ of $\boldsymbol{\varphi}_t(z)$; the probability vector of $\{Z_t \in U_{x-1}\}$ is $\boldsymbol{f}_{t-1}C_t$, which is a part of $z^{(x-1)}$ of $\boldsymbol{\varphi}_t(z)$. Given $\{Z_{t-1} \in U_x\}$, the probability generating function of Z_t is $(\boldsymbol{f}_{t-1}(x)z^x)(A_t + zB_t + z^{-1}C_t)$. Hence on condition that A_t, B_t, C_t do not depend on x, we have $\boldsymbol{\varphi}_t(z) = \boldsymbol{\varphi}_{t-1}(z)(A_t + zB_t + z^{-1}C_t)$. (Theorem 2.2 and Theorem 4.2′ can be recognized by a similar reason.)

From this theorem, we have

$$\varphi_n(z) = \boldsymbol{\varphi}_0(z)\prod_{t=1}^{n}(A_t + zB_t + z^{-1}C_t)\mathbf{1}', \tag{4.4}$$

where

$$\boldsymbol{\varphi}_0(z) = \sum_{x=1}^{l_0} \boldsymbol{f}_0(x)z^x = \sum_{x=0}^{l_0} \boldsymbol{\pi}_x z^x.$$

COROLLARY. *If X_n is a Markov chain imbeddable variable of Binomial type (M.V.B.), we have $C_t = \mathbf{0}$, $t = 1, 2, \ldots$, and*

$$\boldsymbol{\varphi}_t(z) = \boldsymbol{\varphi}_{t-1}(z)(A_t + zB_t), \qquad t \geq 0.$$

This is Theorem 2.2 of Koutras and Alexandrou (1995).

It is convenient for evaluating the distribution of run statistics by using Theorem 4.2. Koutras and Alexandrou (1995) discussed the $N_{n,k}$, $M_{n,k}$, $G_{n,k}$, and scan statistics. We will discuss the $E_{n,k}$ runs in Section 5.

Similar to Theorem 2.3, we have

THEOREM 4.3. *If $A_t(x) = A$, $B_t(x) = B$, $C_t(x) = C$, for all $x = 0, 1, \ldots, l_t$, $(t = 0, 1, \ldots, n)$, and $l_0 = 0$, we have*

$$E(X_n) = \boldsymbol{\varphi}_0(1)\sum_{i=1}^{n} D^{i-1}(B - C)\mathbf{1}',$$

$$E(X_n^2) = \boldsymbol{\varphi}_0(1)\sum_{i=1}^{n}\left[\sum_{s=1}^{i-1} D^{s-1}(B - C)D^{i-1-s}(B - C)\right.$$
$$\left. + D^{i-1}(B - C)\sum_{s=1}^{n-i} D^{s-1}(B - C) + D^{i-1}(B + C)\right]\mathbf{1}'.$$

And their generating functions are

$$M_X(w) = \sum_{n=1}^{\infty} E(X_n)w^n = \frac{w}{1-w}\varphi_0(1)(I - wD)^{-1}(B - C)\mathbf{1}',$$

$$M_{X^2}(w) = \sum_{n=1}^{\infty} E(X_n^2)w^n$$

$$= \frac{w}{1-w}\varphi_0(1)(I - wD)^{-1}[2w(B - C)(I - wD)^{-1}(B - C) + (B + C)]\mathbf{1}',$$

where $D = A + B + C$.

Clearly, we can extend these results of this section to random vector. Similarly as the notations and methods in Section 2, we have the following definition and theorems.

DEFINITION 4.1′. The random vector $\boldsymbol{X}_n$ is called a Markov chain imbeddable vector of returnable type (M.V.R.), if

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

(2) there exists a partition $\{\boldsymbol{U}_{\boldsymbol{x}} : \boldsymbol{x} \geq \mathbf{0}\}$ on the state space Ω,

(3) for every $\boldsymbol{x}$,

$$\Pr(\boldsymbol{X}_n = \boldsymbol{x}) = \Pr(Z_n \in \boldsymbol{U}_{\boldsymbol{x}})$$

and (4)

$$\Pr(Z_t \in \boldsymbol{U}_{\boldsymbol{x}+\boldsymbol{x}^\star} \mid Z_{t-1} \in \boldsymbol{U}_{\boldsymbol{x}}) = 0, \quad \text{if} \quad \boldsymbol{x}^\star \neq \mathbf{0}, \quad \text{or} \quad \boldsymbol{x}^\star \neq \pm \boldsymbol{e}_k, \quad (k = 0, 1, \dots, m).$$

Let

$$A_t(\boldsymbol{x}) = (\Pr(Z_t = U_{\boldsymbol{x},j} \mid Z_{t-1} = U_{\boldsymbol{x},i})),$$
$$B_t^{(k)}(\boldsymbol{x}) = (\Pr(z_t = U_{\boldsymbol{x}+\boldsymbol{e}_k,j} \mid Z_{t-1} = U_{\boldsymbol{x},i})),$$
$$C_t^{(k)}(\boldsymbol{x}) = (\Pr(Z_t = U_{\boldsymbol{x}-\boldsymbol{e}_k,j} \mid Z_{t-1} = U_{\boldsymbol{x},i})), \quad (k = 0, 1, \dots, m).$$

We have

THEOREM 4.1′. *The probability vectors* $\boldsymbol{f}_t(\boldsymbol{x})$, $\mathbf{0} \leq \boldsymbol{x} \leq \boldsymbol{l}_t$, $t = 0, 1, \dots, n$, *satisfy*

$$\boldsymbol{f}_t(\boldsymbol{x}) = \boldsymbol{f}_{t-1}(\boldsymbol{x})A_t(\boldsymbol{x}) + \sum_{k=0}^{m} \boldsymbol{f}_{t-1}(\boldsymbol{x} - \boldsymbol{e}_k)B_t^{(k)}(\boldsymbol{x} - \boldsymbol{e}_k)$$
$$+ \sum_{k=0}^{m} \boldsymbol{f}_{t-1}(\boldsymbol{x} + \boldsymbol{e}_k)C_t^{(k)}(\boldsymbol{x} + \boldsymbol{e}_k) \quad (\mathbf{0} \leq \boldsymbol{x} \leq \boldsymbol{l}_t), \quad (t = 1, \dots, n),$$

and

$$\boldsymbol{f}_t(\boldsymbol{x}) = \mathbf{0}, \quad \text{if} \quad \boldsymbol{x} \notin \{\boldsymbol{x} : \mathbf{0} \leq \boldsymbol{x} \leq \boldsymbol{l}_t\}.$$

The probability distribution function of a M.V.R. $\boldsymbol{X}_n$ *is given by*

$$\Pr(\boldsymbol{X}_n = \boldsymbol{x}) = \boldsymbol{f}_n(\boldsymbol{x})\mathbf{1}'. \tag{4.5}$$

THEOREM 4.2′. *If* $A_t(\boldsymbol{x}) = A_t$, $B_t^{(k)}(\boldsymbol{x}) = B_t^{(k)}$, *and* $C_t^{(k)}(\boldsymbol{x}) = C_t^{(k)}$ *for all* $\boldsymbol{x}$, $(k = 0, 1, \ldots, m)$, *we have*

$$\varphi_t(z) = \varphi_{t-1}(z)\left(A_t + \sum_{k=0}^{m} z_k B_t^{(k)} + \sum_{k=0}^{m} z_k^{-1} C_t^{(k)}\right), \qquad t = 1, 2, \ldots. \tag{4.6}$$

Using these theorems, we obtain the joint distribution of $(E_{n,k_i}^{(i)}, i = 0, 1, \ldots, m)$ in the next section.

5. Distributions of $E_{n,k}$ and $(E_{n,k_i}^{(i)},\ i = 0, 1, \ldots, m)$

5.1 *Success runs of length exactly* k, $E_{n,k}$

Let $E_{t,k}$ be the number of success runs of size exactly k until the t-th trial, in a sequence of Bernoulli trials. We regard the value 0 as success and the value 1 as failure.

Let us denote by x the number of "0" runs of length exactly k up to the t-th trial. Let $y^\star$ be the number of trailing "0" (i.e. the number of last consecutive "0" counting backwards, $y^\star = 0$ if the t-th trial is "1").

Let $l_n = [\frac{n+1}{k+1}]$, and

$$y = \begin{cases} y^* & \text{if } \; y^* \le k-1 \\ -1 & \text{if } \; y^* > k \\ -2 & \text{if } \; y^* = k \end{cases}.$$

Then we can give a Markov chain $\{Z_t = (x, y) : t = 0, 1, \ldots\}$. The state space is

$$\Omega = \{(x, y) \mid 0 \le x \le l_n, -2 \le y \le k-1\} = \bigcup_{x=0}^{l_n} U_x,$$

where $U_x = \{(x, y) \mid -2 \le y \le k-1\} = \{(x, 0), (x, 1), \ldots, (x, k-1), (x, -2), (x, -1)\}$, and $s = k + 2$.

The state $(0, -2)$ is an additional hypothetical state (inaccessible), in order to make the cardinality of U_0 equal to $|U_x| = k + 2$. Because the additional state $(0, -2)$, we have $\boldsymbol{f}_{t-1}(0)C_t = \mathbf{0}$. And the Markov chain can not be transferred to a negative state.

With this set up, the random variable $E_{n,k}$ satisfies Definition 4.1, with

$$\Pr(Z_t = (x, y+1) \mid Z_{t-1} = (x, y)) = p_t, \qquad 0 \le y \le k-2, \tag{5.1}$$

$$\Pr(Z_t = (x, -1) \mid Z_{t-1} = (x, -1)) = p_t, \tag{5.2}$$

$$\Pr(Z_t = (x, 0) \mid Z_{t-1} = (x, y)) = 1 - p_t, \qquad -2 \le y \le k-2, \tag{5.3}$$

$$\Pr(Z_t = (x+1, -2) \mid Z_{t-1} = (x, k-1)) = p_t, \tag{5.4}$$

$$\Pr(Z_t = (x-1, -1) \mid Z_{t-1} = (x, -2)) = p_t. \tag{5.5}$$

The transition matrices A_t, B_t and C_t can be obtained by these formulas.

As an illustration, we consider $E_{5,2}$. In the example, $n = 5$, $k = 2$ and $l_5 = 2$, and $U_x = \{(x,0),(x,1),(x,-2),(x,-1)\}$, $(x = 0,1,2)$, and $s = 4$. We have

$$A_t = \begin{pmatrix} (x,0) & (x,1) & (x,-2) & (x,-1) \\ q_t & p_t & 0 & 0 \\ q_t & 0 & 0 & 0 \\ q_t & 0 & 0 & 0 \\ q_t & 0 & 0 & p_t \end{pmatrix},$$

$$B_t = \begin{pmatrix} (x,0) & (x,1) & (x,-2) & (x,-1) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & p_t & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$C_t = \begin{pmatrix} (x,0) & (x,1) & (x,-2) & (x,-1) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_t \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The initial probabilities are $\boldsymbol{\pi_0} = (1,0,0,0)$, and $\boldsymbol{\pi_x} = \mathbf{0}$, for all $x \neq 0$. Using Theorem 4.2, $\boldsymbol{\varphi}_0(z) = (1,0,0,0)$, $\boldsymbol{\varphi}_1(z) = (q_1,p_1,0,0)$, $\boldsymbol{\varphi}_2(z) = (q_1q_2 + p_1q_2, q_1p_2, zp_1p_2, 0), \ldots$.

For a sequence of Bernoulli trials of length $n = 5$, in which the t-th trial's probability of success is $p_t = \frac{1}{(t+1)}$ $(t = 1,2,\ldots,5)$, we have $\varphi_5(z) = \boldsymbol{\varphi}_5(z)\mathbf{1}' = 0.7931 + 0.2028z + 0.0042z^2$. The result agrees with Fu and Koutras (1994).

5.2 *Runs of length exactly* k_i, $(E^{(0)}_{n,k_0}, \ldots, E^{(m)}_{n,k_m})$

Let $l^{(i)}_n = [\frac{n+1}{k_i+1}]$, and

$$y_i = \begin{cases} y_i^* & \text{if } \ y_i^* \leq k_i - 1 \\ -1 & \text{if } \ y_i^* > k_i \\ -2 & \text{if } \ y_i^* = k_i \end{cases}.$$

The state space is

$$\begin{aligned} \Omega &= \{(\boldsymbol{x}; y_0, y_1, \ldots, y_m) \mid \mathbf{0} \leq \boldsymbol{x} \leq \boldsymbol{l}_n, -2 \leq y_i \leq k_i - 1, i = 0,1,\ldots,m\} \\ &= \bigcup_{\mathbf{0} \leq \boldsymbol{x} \leq \boldsymbol{l}_n} \boldsymbol{U_x}, \end{aligned}$$

where $\boldsymbol{U_x} = \{(\boldsymbol{x}; y_0, y_1, \ldots, y_m) \mid -2 \leq y_i \leq k_i - 1, i = 0,1,\ldots,m\}$, and $s = \sum_{i=0}^{m}(k_i + 1) + 1$. And we have

$$\Pr(Z_t = (\boldsymbol{x}; 0, \ldots, y_i + 1, 0, \ldots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \ldots, y_i, 0, \ldots, 0)) = p_t^{(i)}$$
$$0 \leq y_i \leq k_i - 2, \qquad 0 \leq i \leq m,$$

$$\Pr(Z_t = (\boldsymbol{x}; 0, \ldots, \overbrace{-1}^{(i)}, \ldots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \ldots, \overbrace{-1}^{(i)}, \ldots, 0)) = p_t^{(i)} \qquad 0 \le i \le m,$$

$$\Pr(Z_t = (\boldsymbol{x}; 0, \ldots, \overbrace{1}^{(j)}, \ldots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \ldots, y_i, \ldots, 0)) = p_t^{(j)} \qquad -2 \le y_i \le k_i - 2, \quad 0 \le i, j \le m, \quad i \ne j$$

$$\Pr(Z_t = (\boldsymbol{x} + \boldsymbol{e}_i; 0, \ldots, \overbrace{-2}^{(i)}, \ldots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \ldots, k_i - 1, 0, \ldots, 0)) = p_t^{(i)} \qquad 0 \le i \le m,$$

and

$$\Pr(Z_t = (\boldsymbol{x} - \boldsymbol{e}_i; 0, \ldots, \overbrace{-1}^{(i)}, \ldots, 0) \mid Z_{t-1} = (\boldsymbol{x}; 0, \ldots, \overbrace{-2}^{(i)}, \ldots, 0)) = p_t^{(i)} \qquad 0 \le i \le m.$$

As an illustration, we consider the distribution of $(E_{5,2}^{(0)}, E_{5,2}^{(1)}, E_{5,3}^{(2)})$ in a sequence of three-state trials of length $n = 5$. For outcomes "0", "1", "2", we assume that the t-th trial's probability is $p_t^{(0)} = \frac{1}{t+1}$, $p_t^{(1)} = \frac{1}{t+2}$, $p_t^{(2)} = 1 - p_t^{(0)} - p_t^{(1)}$.

Using Theroem 4.3′, we get

$$A_t + \sum_{k=0}^{2} z_k B_t^{(k)} + \sum_{k=0}^{2} z_k^{-1} C_t^{(k)} =$$

$$\begin{pmatrix}
(000) & 0 & p_t^{(2)} & 0 & 0 & 0 & p_t^{(1)} & 0 & 0 & p_t^{(0)} & 0 & 0 \\
(001) & 0 & 0 & p_t^{(2)} & 0 & 0 & p_t^{(1)} & 0 & 0 & p_t^{(0)} & 0 & 0 \\
(002) & 0 & 0 & 0 & z_2 p_t^{(2)} & 0 & p_t^{(1)} & 0 & 0 & p_t^{(0)} & 0 & 0 \\
(00-2) & 0 & 0 & 0 & 0 & z_2^{-1} p_t^{(2)} & p_t^{(1)} & 0 & 0 & p_t^{(0)} & 0 & 0 \\
(00-1) & 0 & 0 & 0 & 0 & p_t^{(2)} & p_t^{(1)} & 0 & 0 & p_t^{(0)} & 0 & 0 \\
(010) & 0 & p_t^{(2)} & 0 & 0 & 0 & 0 & z_1 p_t^{(1)} & 0 & p_t^{(0)} & 0 & 0 \\
(0-20) & 0 & p_t^{(2)} & 0 & 0 & 0 & 0 & 0 & z_1^{-1} p_t^{(1)} & p_t^{(0)} & 0 & 0 \\
(0-10) & 0 & p_t^{(2)} & 0 & 0 & 0 & 0 & 0 & p_t^{(1)} & p_t^{(0)} & 0 & 0 \\
(100) & 0 & p_t^{(2)} & 0 & 0 & 0 & p_t^{(1)} & 0 & 0 & 0 & z_0 p_t^{(0)} & 0 \\
(-200) & 0 & p_t^{(2)} & 0 & 0 & 0 & p_t^{(1)} & 0 & 0 & 0 & 0 & z_0^{-1} p_t^{(0)} \\
(-100) & 0 & p_t^{(2)} & 0 & 0 & 0 & p_t^{(1)} & 0 & 0 & 0 & 0 & p_t^{(0)}
\end{pmatrix},$$

and $\boldsymbol{\varphi}_0(z_0, z_1, z_2) = \sum_{0 \le \boldsymbol{x} \le l_0} \boldsymbol{\pi}_{\boldsymbol{x}} z_0^{x_0} z_1^{x_1} z_2^{x_2} = \boldsymbol{\pi}_{\mathbf{0}} = (1, 0, \ldots, 0)$, where A_t and $B_t^{(k)}, C_t^{(k)}$, $(k = 0, 1, 2)$ can be given from the above formulas. Hence we have $\varphi_5(z_0, z_1, z_2) = \varphi_0(z_0, z_1, z_2) \Pi_{t=1}^{5} (A_t + \sum_{k=0}^{2} z_k B_t^{(k)} + \sum_{k=0}^{2} z_k^{(-1)} C_t^{(k)}) \mathbf{1}' = 0.04167 z_0^2 + 0.014047 z_0 z_1 + 0.041359 z_0 z_2 + 0.147370 z_0 + 0.001587 z_1^2 + 0.020952 z_1 z_2 + 0.096744 z_1 + 0.129430 z_2 + 0.544336$. Hence Table 2 gives the exact joint distribution of $(E_{5,2}^{(0)}, E_{5,2}^{(1)}, E_{5,3}^{(2)})$.

Let $z_1 = 1$ and $z_2 = 1$, we have $\varphi_5(z_0, z_1 = 1, z_2 = 1) = 0.004167 z_0^2 + 0.202776 z_0 + 0.793049$. This is the example of Subsection 5.1.

Table 2. The exact joint distribution of $(E^{(0)}_{5,2}, E^{(1)}_{5,2}, E^{(2)}_{5,2})$.

		$X_1 = 0$	$X_1 = 1$	$X_1 = 2$	
$X_0 = 0$	$X_2 = 0$	0.544336	0.096744	0.001587	
	$X_2 = 1$	0.129430	0.020952		0.793049
$X_0 = 1$	$X_2 = 0$	0.147370	0.014047		
	$X_2 = 1$	0.041359			0.202776
$X_0 = 2$	$X_2 = 0$	0.004167			
	$X_2 = 1$				0.004167

6. Waiting time problems

In Sections 3 and 5, we established some proper Markov chains $\{Z_t : t \geq 0\}$. Actually, with these set up, these run statistics X_n $(n = 1, 2, \ldots)$ can be imbedded into the same Markov chain. Because this reason, we can also use our approach to find the waiting time distribution.

For convenience, we only consider the sequence with outcomes "0" and "1", i.e. $m = 1$ case. Let E_{x_0} (F_{x_1}) denote the event that x_0 (x_1) "0"-runs ("1"-runs) of length k_0 (k_1) occur. We discuss the waiting time problems of E_{x_0} and F_{x_1} (Uchida and Aki (1995)).

For example, we consider the sooner waiting time W_S of E_{x_0} and F_{x_1} in nonoverlapping run case. Because

$$\begin{aligned}\{W_S = t\} = &\bigcup_{j=0}^{x_1-1} [\{Z_{t-1} = ((x_0 - 1, j); (k_0 - 1, 0))\} \cap \{Z_t = ((x_0, j); (0, 0))\}] \\ &\cup \bigcup_{j=0}^{x_0-1} [\{Z_{t-1} = ((j, x_1 - 1); (0, k_1 - 1))\} \cap \{Z_t = ((j, x_1); (0, 0))\}],\end{aligned}$$

we can obtain the probability function of the sooner waiting time W_S,

$$\begin{aligned}\Pr(W_S = t) = &\sum_{j=0}^{x_1-1} \boldsymbol{f}_{t-1}((x_0 - 1, j))p_0 \cdot (0, 0, \ldots, 0, 1)' \\ &+ \sum_{j=0}^{x_0-1} \boldsymbol{f}_{t-1}((j, x_1 - 1))p_1 \cdot (0, \ldots, \overbrace{1}^{(k_1)}, \ldots, 0)'.\end{aligned}$$

The probability generating function of W_S is

$$\begin{aligned}\psi_{x_0,x_1}(\omega) &= \sum_{t=0}^{\infty} \Pr(W_S = t) \cdot \omega^t \\ &= \sum_{j=0}^{x_1-1} \sum_{t=0}^{\infty} \boldsymbol{f}_{t-1}((x_0 - 1, j))\omega^{t-1} \cdot (0, \ldots, 1)' \cdot (\omega p_0)\end{aligned}$$

$$+\sum_{j=0}^{x_0-1}\sum_{t=0}^{\infty} f_{t-1}((j,x_1-1))\omega^{t-1}\cdot(0,\dots,\overbrace{1}^{(k_1)},\dots,0)'\cdot(\omega p_1).$$

First, we consider the function

$$\begin{aligned}\Psi(z_0,z_1,\omega) &= \sum_{x_0,x_1}\sum_{t=0}^{\infty} f_t(x_0,x_1)z_0^{x_0}z_1^{x_1}\omega^t = \sum_{t=0}^{\infty}\varphi_t(z_0,z_1)\omega^t\\ &= \varphi_0(z_0,z_1)\cdot\sum_{t=0}^{\infty}[A_t + z_0B_t^{(0)} + z_1B_t^{(1)}]^t\omega^t,\end{aligned}$$

when i.i.d. case,

$$\begin{aligned}\Psi(z_0,z_1,\omega) &= \Phi(z_0,z_1,\omega)\\ &= (1,0,\dots,0)\cdot[I-\omega(A+z_0B^{(0)}+z_1B^{(1)})]^{-1}\\ &= \Bigg(\frac{1-(a_0-1)(a_1-1)}{\Delta},\frac{a_0(\omega p_1)}{\Delta},\dots,\frac{a_0(\omega p_1)^{k_1-1}}{\Delta},\\ &\qquad\qquad \frac{a_1(\omega p_0)}{\Delta},\dots,\frac{a_1(\omega p_0)^{k_0-1}}{\Delta}\Bigg),\end{aligned}$$

where $a_0 = \sum_{i=0}^{k_0-1}(\omega p_0)^i$, $a_1 = \sum_{i=0}^{k_1-1}(\omega p_1)^i$ and $\Delta = 1-(a_0-1)(a_1-1) - z_0a_1(\omega p_0)^{k_0} - z_1a_0(\omega p_1)^{k_1}$.

So, we have

$$\begin{aligned}\psi_{x_0,x_1}(\omega) &= (\omega p_0)\sum_{j=0}^{x_1-1}\text{ the coefficient of } z_0^{x_0-1}z_1^j \text{ of } \frac{a_1(\omega p_0)^{k_0-1}}{\Delta}\\ &\quad + (\omega p_1)\sum_{j=0}^{x_0-1}\text{ the coefficient of } z_0^{j}z_1^{x_1-1} \text{ of } \frac{a_0(\omega p_1)^{k_1-1}}{\Delta}.\end{aligned}$$

THEOREM 6.1. *The prabability generating function of the sooner waiting time W_S is*

$$\begin{aligned}\psi_{x_0,x_1}(\omega) = &\left[\frac{a_1(\omega p_0)^{k_0}}{1-(a_0-1)(a_1-1)}\right]^{x_0}\sum_{j=0}^{x_1-1}\binom{x_0-1+j}{x_0-1}\left[\frac{a_0(\omega p_1)^{k_1}}{1-(a_0-1)(a_1-1)}\right]^j\\ &+\left[\frac{a_0(\omega p_1)^{k_1}}{1-(a_0-1)(a_1-1)}\right]^{x_1}\\ &\cdot\sum_{j=0}^{x_0-1}\binom{x1-1+j}{x_1-1}\left[\frac{a_1(\omega p_0)^{k_0}}{1-(a_0-1)(a_1-1)}\right]^j.\end{aligned}$$

Similarly, we can discuss the Markov chain dependent trial case. We can also obtain the distribution and the probability generating function of the later waiting time W_L, because of

$$\{W_L = t\} = \bigcup_{j=x_1}^{\infty} [\{Z_{t-1} = ((x_0 - 1, j); (k_0 - 1, 0))\} \cap \{Z_t = ((x_0, j); (0, 0))\}]$$
$$\cup \bigcup_{j=x_0}^{\infty} [\{Z_{t-1} = ((j, x_1 - 1); (0, k_1 - 1))\} \cap \{Z_t = ((j, x_1); (0, 0))\}].$$

We can discuss the M-run and the G-run, because we have
(1) overlapping run,

$$\{W_S = t\} = \bigcup_{j=0}^{x_1-1} \{Z_t = ((x_0, j); (-1, 0))\} \cup \bigcup_{j=0}^{x_0-1} \{Z_t = ((j, x_1); (0, -1))\},$$
$$\{W_L = t\} = \bigcup_{j=x_1}^{\infty} \{Z_t = ((x_0, j); (-1, 0))\} \cup \bigcup_{j=x_0}^{\infty} \{Z_t = ((j, x_1); (0, -1))\}.$$

(2) run of length k or more,

$$\{W_S = t\} = \bigcup_{j=0}^{x_1-1} [\{Z_t = ((x_0, j); (-1, 0))\} \cap \{Z_{t-1} = ((x_0 - 1, j); (k_0 - 1, 0))\}]$$
$$\cup \bigcup_{j=0}^{x_0-1} [\{Z_t = ((j, x_1); (0, -1))\} \cap \{Z_{t-1} = ((j, x_1 - 1); (0, k_1 - 1))\}],$$
$$\{W_L = t\} = \bigcup_{j=x_1}^{\infty} [\{Z_t = ((x_0, j); (-1, 0))\} \cap \{Z_{t-1} = ((x_0 - 1, j); (k_0 - 1, 0))\}]$$
$$\cup \bigcup_{j=x_0}^{\infty} [\{Z_t = ((j, x_1); (0, -1))\} \cap \{Z_{t-1} = ((j, x_1 - 1); (0, k_1 - 1))\}].$$

Acknowledgements

The authors would like to thank the referees for their valuable comments.

References

Aki, S. (1985). Discrete distributions of order k on a binary sequence, *Ann. Inst. Statist. Math.*, **37**, 205–224.

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. (1988). Some characteristics of the binomial distribution of order k and related distributions, *Statistical Theory and Data Analysis II*; *Proceedings of the 2nd Pacific*

Area Statistical Conference (ed. K. Matusita), 211–222, Elsevier Science Publishers B.V., North-Holland.

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 3rd Pacific Area Statistical Conference* (eds. K. Matusita, M. L. Puri and T. Hayakawa), 467–474, VSP International Science Publishers, Zeist.

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, *Statist. 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.

Hirano, K. (1986). Some properties of the distributions of order k, *Fibonacci Numbers and Their Applications* (eds. A. N. Philippou, A. F. Horadam and G. E. Bergum), 43–53, Reidel, Dordrecht.

Hirano, K. and Aki, S. (1993). On number of occurrences of success runs of specified length in a two-state Markov chain, *Statist. Sinica*, **3**, 313–320.

Hirano, K., Aki, S., Kashiwagi, N. and Kuboki, H. (1991). On Ling's binomial and negative binomial distributions of order k, *Statist. Probab. Lett.*, **11**, 503–509.

Koutras, M. V. and Alexandrou, V. A. (1995). Runs, scans and urn model distributions: A unified Markov chain approach, *Ann. Inst. Statist. Math.*, **47**, 743–766.

Koutras, M. V. and Alexandrou, V. A. (1997). Sooner waiting time problems in a sequence of trinomial trials, *J. Appl. Probab.*, **34**, 593–609.

Ling, K. D. (1988). On binomial distributions of order k, *Statist. Probab. Lett.*, **6**, 247–250.

Ling, K. D. (1989). A new class of negative binomial distributions of order k, *Statist. Probab. Lett.*, **7**, 371–376.

Ling, K. D. and Low, T. (1993). On the soonest and latest waiting time distributions: Succession quotas, *Comm. Statist. Theory Methods*, **22**(8), 2207–2221.

Mood, A. M. (1940). The distribution theory of runs, *Ann. Math. Siatist.*, **11**, 367–392.

Philippou, A. N. (1986). Distributions and Fibonacci polynomials of order k, longest runs, and reliability of consecutive-k-out-of-n: F systems, *Fibonacci Numbers and Their Applications* (eds. A. N. Philippou, G. E. Bergum and A. F. Horadam), 203–227, Reidel, Dordrecht.

Philippou, A. N. and Makri, F. S. (1986). Successes, runs and longest runs, *Statist. Probab. Lett.*, **4**, 211–215.

Philippou, A. N., Antzoulakos, D. L. and Tripsiannis, G. A. (1988). Multivariate distributions of order k, *Statist. Probab. Lett.*, **7**, 207–216.

Schwager, S. J. (1983). Run probabilities in sequences of Markov-Dependent trials, *J. Amer. Statist. Assoc.*, **78**, 168–175.

Uchida, M. and Aki, S. (1995). Sooner and later waiting time problems in a two-state Markov chain, *Ann. Inst. Statist. Math.*, **47**, 415–433.