Appl. Math. J. Chinese Univ.
2018, 33(1): 25-34

Waiting times and stopping probabilities for patterns in Markov chains

ZHAO Min-zhi[1] XU Dong[1] ZHANG Hui-zeng[2,*]

Abstract. Suppose that $\mathcal{C}$ is a finite collection of patterns. Observe a Markov chain until one of the patterns in $\mathcal{C}$ occurs as a run. This time is denoted by τ. In this paper, we aim to give an easy way to calculate the mean waiting time $E(\tau)$ and the stopping probabilities $P(\tau = \tau_A)$ with $A \in \mathcal{C}$, where τ_A is the waiting time until the pattern A appears as a run.

§1 Introduction

Suppose that $\{Z_n\}_{n\geq 1}$ is a discrete time homogenous Markov chain with finite state space Δ. A finite sequence of elements from Δ is called a pattern. We will use a capital letter to denote a pattern. Use $\mathcal{C}$ to denote a finite collection of patterns. For example, if $\Delta = \{0, 1\}$, then $A = 1011$ is a pattern while $\mathcal{C} = \{101, 11\}$ is a finite collection of patterns. For a pattern A, use τ_A to denote the waiting time until A occurs as a run in the sequence $Z_1, Z_2, \cdots$. Let $\tau = \tau_{\mathcal{C}} = \min\{\tau_A : A \in \mathcal{C}\}$ be the waiting time till one of the patterns appears. We are interested in the calculation of $E(\tau)$ and $P(\tau = \tau_A)$ with $A \in \mathcal{C}$.

In many applications, such as quality control, hypothesis testing, reliability theory and scan statistics, the distribution of τ is very important. Naus [9,10] used a window with length w to scan a process until time T and then got a scan statistic. The distribution of this scan statistic can be transformed into the distribution of $\tau_{\mathcal{C}}$ with some special collection of patterns. For example, if $\Delta = \{0, 1\}$, $w = 4$ and the scan statistic is

$$S_T = \max_{1\leq i\leq T-3}(Z_i + Z_{i+1} + Z_{i+2} + Z_{i+3}),$$

then S_T denotes the maximal number of 1 appears in a window of length 4 until time T. In this case, $P(S_T \geq 2) = P(\tau_{\mathcal{C}} \leq T)$, where $\mathcal{C} = \{11, 101, 1001\}$. Another interesting application is Penney-Ante game which is developed by Walter Penney (see [11]). It is a game with two

Received: 2017-01-13. Revised: 2018-01-16.
MR Subject Classification: 60J10, 60J22.
Keywords: pattern, Markov chain, stopping probability, waiting time.
Digital Object Identifier(DOI): https://doi.org/10.1007/s11766-018-3522-z.
Supported by the National Natural Science Foundation of China (11771286, 11371317) and by the Zhejiang Provincial Natural Science Foundation of China (LQ18A010007).
* Corresponding author.

players. Player I chooses a triplet of outcomes namely A. Then payer II chooses a different triplet namely B. An unbiased coin is flipped repeatedly until A or B is observed. If A occurs first, then player I wins the game. Otherwise player II wins. Clearly, the wining probability for player II is $P(\tau_{\mathcal{C}} = \tau_B)$, where $\mathcal{C} = \{A, B\}$. After player I has selected A, the most important thing for player II is to find an optimal strategy, that is he should find a triplet B that maximizes his winning probability. In fact, such an optimal strategy exists (see [1]).

Thanks to its importance, the occurrence of patterns has been studied by many people. When $Z_1, Z_2, \cdots$ are i.i.d., Li [8], Gerber and Li [5] used the Martingale method to study the problem. Later in 1981, Guibas and Odlyzko [7] used the combinatorial method to obtain the linear equations of $E(\tau)$ and $P(\tau = \tau_A)$. When $\{Z_n\}$ is a Markov chain, in 1990, Chrysaphinou and Papastavridis [2] used the combinatorial method to obtain the linear equations of $E(\tau)$. In 2002, Fu and Chang [3] studied $E(\tau)$ by using Markov chain embedding method. Later Glaz, Kulldorff and etc. [6], Pozdnyakov [12] introduced gambling teams and used Martingale theory to study $E(\tau)$. In 2014, Gava and Salotti [4] obtained the system of linear equations of $P(\tau = \tau_A)$ with $A \in \mathcal{C}$ based on the results of [6] and [12].

When $\{Z_n\}$ is a Markov chain, though the mean waiting time $E(\tau)$ and the stopping probabilities $P(\tau = \tau_A)$ were obtained in [4], [6] and [12], the method is complicated. Briefly speaking, the method is divided into four steps. Firstly, define the sets $\mathcal{D}' = \{lA : l \in \Delta, A \in \mathcal{C}\}$ and $\mathcal{C}' = \{lmA : l, m \in \Delta, A \in \mathcal{C}\}$. Use $\mathcal{D}''$ and $\mathcal{C}''$ to denote the collection of patterns excluding from $\mathcal{D}'$ and from $\mathcal{C}'$, respectively, the patterns that cannot occur at time τ. Set $K' = |\mathcal{C}| + |\mathcal{D}''|$ and $M' = |\mathcal{C}''|$. Secondly, introduce the gambling teams, compute the profit matrix W that has $(K' + M')M'$ elements, and compute the probability of occurrence of the i-th ending scenario with $i = 1, 2, \cdots, K' + M'$. Thirdly, solve a linear system of M' equations in M' variables and then obtain the mean waiting time $E(\tau)$. Finally, solve about M' linear systems involving M' equations and M' variables and then get the stopping probabilities $P(\tau = \tau_A)$.

In this paper, we aim to find an easy and effective method to calculate $E(\tau)$ and $P(\tau = \tau_A)$. Inspired by the paper [7], we use the combinative probabilistic analysis and the Markov property. The main result of our paper is Theorem 2.1. It extend Theorem 3.3 of [7] to Markov case. Corollary 2.3 gives a better way to obtain $E(\tau)$ and $P(\tau = \tau_A)$ with $A \in \mathcal{C}$: solving only a single linear system involving $|\Delta| + |\mathcal{C}|$ equations and $|\Delta| + |\mathcal{C}|$ variables. The rest of the paper is organized as follows. In §2, the main results and the proofs are given. In §3, some examples are discussed.

§2 Main results

In our paper, suppose that $\{Z_n\}_{n\geq 1}$ is a discrete time homogenous Markov chain with finite state space Δ, initial distribution $\mu_i = P(Z_1 = i)$ and one-step transition probability $P_{ij} = P(Z_{n+1} = j|Z_n = i)$. We will make the following three assumptions.

(A.1) No pattern in $\mathcal{C}$ is a subpattern of another pattern in $\mathcal{C}$.

(A.2) For any $K = K_1K_2 \cdots K_m \in \mathcal{C}$, $P_{K_1K_2} \cdots P_{K_{m-1}K_m} > 0$.

(A.3) That $P(\tau < \infty) = 1$ and $E(\tau) < \infty$.

For a pattern K, let K_i denote the i-th element of K, $|K|$ denote the length of K, that is, $K = K_1 K_2 \cdots K_{|K|}$. Let $X_K^{(j)} = I_{\{j\}}(K_{|K|})$. For patterns $K = K_1 \cdots K_s$ and $T = T_1 \cdots T_t$, let $\{KT\}$ be a subset of $\{1, 2, \cdots, s \wedge t\}$ such that an integer k is in $\{KT\}$ if and only if $K_{s-k+1} \cdots K_s = T_1 \cdots T_k$. Note that in [7], the correlation of K and T, denoted by $\mathrm{cor}(K, T)$, is defined as a string over $\{0, 1\}$ with the same length as K. The k-th bit (from the right) of $\mathrm{cor}(K, T)$ is 1 if and only if $k \in \{KT\}$. For example, if $K = 101001$ and $T = 10010$, then $\mathrm{cor}(K, T) = 001001$ and $\{KT\} = \{1, 4\}$. Here the correlation between two patterns is different to the traditional correlation between two random variables.

For any $i \in \Delta$ and any pattern K, let

$$P_{i \to K} = P\left((Z_2, \cdots, Z_{|K|+1}) = K | Z_1 = i\right) = P_{iK_1} P_{K_1 K_2} \cdots P_{K_{|K|-1} K_{|K|}}.$$

The pattern of length 0 is denoted by ϕ. Set $P_{i \to \phi} = 1$. For any pattern $K, T \in \mathcal{C}$, let

$$\tilde{g}_{KT}(z) = \begin{cases} \displaystyle\sum_{\substack{r \in \{KT\} \\ 1 \le r < |T|}} z^r \cdot P_{T_r \to T_{r+1} \cdots T_{|T|}} \Bigg/ P_{T_1 \to T_2 \cdots T_{|T|}}, & K \neq T, \\ \left(\displaystyle\sum_{\substack{r \in \{KT\} \\ 1 \le r < |T|}} z^r \cdot P_{T_r \to T_{r+1} \cdots T_{|T|}} + z^{|T|} \right) \Bigg/ P_{T_1 \to T_2 \cdots T_{|T|}}, & K = T. \end{cases}$$

For $i \in \Delta$, $K \in \mathcal{C}$ and $n \ge 1$, define

$$S_i(n) = P(Z_n = i, \tau > n) \text{ and } S_K(n) = P(\tau = \tau_K = n).$$

Now, define the corresponding generating functions

$$F_i(z) = \sum_{n=1}^{\infty} S_i(n) \cdot z^{-n} \text{ and } f_K(z) = \sum_{n=1}^{\infty} S_K(n) \cdot z^{-n},$$

where $z \ge 1$. Our main result is the following Theorem.

Theorem 2.1. *For any $z \ge 1$, the functions $F_i(z)$ and $f_K(z)$ with $i \in \Delta$ and $K \in \mathcal{C}$ satisfy the following system of linear equations:*

$$\begin{cases} \sum\limits_{i \in \Delta} F_i(z) \cdot P_{ij} = z \cdot F_j(z) + z \cdot \sum\limits_{K \in \mathcal{C}} f_K(z) \cdot X_K^{(j)} - \mu_j, & j \in \Delta, \\ \sum\limits_{i \in \Delta} F_i(z) \cdot P_{iT_1} = \sum\limits_{K \in \mathcal{C}} f_K(z) \cdot \tilde{g}_{KT}(z) - \mu_{T_1}, & T \in \mathcal{C}. \end{cases} \tag{2.1}$$

Proof. Firstly, for $j \in \Delta$ and $n \ge 1$,

$$\begin{aligned} \sum_{i \in \Delta} S_i(n) \cdot P_{ij} &= P(\tau > n, Z_{n+1} = j) \\ &= P(\tau > n+1, Z_{n+1} = j) + \sum_{K \in \mathcal{C}} P(\tau = \tau_K = n+1, Z_{n+1} = j) \\ &= S_j(n+1) + \sum_{K \in \mathcal{C}} S_K(n+1) \cdot X_K^{(j)}. \end{aligned}$$

Thus we have,

$$\sum_{n=1}^{\infty} \sum_{i \in \Delta} S_i(n) \cdot z^{-n} \cdot P_{ij} = z \cdot \sum_{n=1}^{\infty} S_j(n+1) \cdot z^{-n-1} + z \cdot \sum_{n=1}^{\infty} \sum_{K \in \mathcal{C}} S_K(n+1) \cdot z^{-n-1} \cdot X_K^{(j)}.$$

Note that

$$S_j(1) + \sum_{K \in \mathcal{C}} S_K(1) \cdot X_K^{(j)} = P(Z_1 = j) = \mu_j.$$

It follows that

$$\sum_{i \in \Delta} F_i(z) \cdot P_{ij} = z \cdot F_j(z) + z \cdot \sum_{K \in \mathcal{C}} f_K(z) \cdot X_K^{(j)} - \mu_j. \tag{2.2}$$

Secondly, for $T \in \mathcal{C}$ and $i \in \Delta$, define

$$S_{i,T}(n) = \begin{cases} 0, & n \leq |T|, \\ P(\tau = \tau_T = n, Z_{n-|T|} = i), & n \geq |T| + 1. \end{cases}$$

Define the corresponding generating function $f_{i,T}(z)$ on $z \geq 1$ as

$$f_{i,T}(z) = \sum_{n=1}^{\infty} S_{i,T}(n) \cdot z^{-n}.$$

Clearly, when $n \geq |T| + 1$, $S_T(n) = \sum_{i \in \Delta} S_{i,T}(n)$. It implies that

$$\sum_{|T|+1}^{\infty} S_T(n) \cdot z^{-n} = \sum_{i \in \Delta} \sum_{|T|+1}^{\infty} S_{i,T}(n) \cdot z^{-n}.$$

Set $P_T = P\left((Z_1, \cdots, Z_{|T|}) = T\right) = \mu_{T_1} \cdot P_{T_1 \to T_2 \cdots T_{|T|}}$. Then we have

$$f_T(z) - z^{-|T|} \cdot P_T = \sum_{i \in \Delta} f_{i,T}(z). \tag{2.3}$$

Thirdly, for $T \in \mathcal{C}$, $i \in \Delta$ and $n \geq 1$,

$$\begin{aligned} S_i(n) \cdot P_{i \to T} &= P\left(\tau > n, Z_n = i, (Z_{n+1}, \cdots, Z_{n+|T|}) = T\right) \\ &= \sum_{r=1}^{|T|} P\left(\tau = n + r, Z_n = i, (Z_{n+1}, \cdots, Z_{n+|T|}) = T\right) \\ &= \sum_{1 \leq r < |T|} \sum_{K \in \mathcal{C}} P\left(\tau = \tau_K = n + r, Z_n = i, (Z_{n+1}, \cdots, Z_{n+|T|}) = T\right) \\ &\quad + P(\tau = \tau_T = n + |T|, Z_n = i). \end{aligned} \tag{2.4}$$

Obviously,

$$P(\tau = \tau_T = n + |T|, Z_n = i) = S_{i,T}(n + |T|). \tag{2.5}$$

For $1 \leq r < |T|$ and $K \in \mathcal{C}$, under the condition that $\tau = \tau_K = n+r$, we have $(Z_{n+r-|K|+1}, \cdots, Z_{n+r}) = K$. If in addition $Z_n = i$ and $(Z_{n+1}, \cdots, Z_{n+|T|}) = T$, then for the reason that K is not a subpattern of T (except that K may be equal to T), we have $|K| \geq r+1$, $K_{|K|-r+1} \cdots K_{|K|} = T_1 \cdots T_r$ and $K_{|K|-r} = i$, that is, $r \in \{KT\}$ and $K_{|K|-r} = i$. Therefore

$$\begin{aligned} &P\left(\tau = \tau_K = n + r, Z_n = i, (Z_{n+1}, \cdots, Z_{n+|T|}) = T\right) \\ =&P\left(\tau = \tau_K = n + r, (Z_{n+r+1}, \cdots, Z_{n+|T|}) = (T_{r+1}, \cdots, T_{|T|})\right) \\ &\cdot I_{\{KT\}}(r) \cdot I_{\{i\}}(K_{|K|-r}) \\ =&S_K(n + r) \cdot P_{T_r \to T_{r+1} \cdots T_{|T|}} \cdot I_{\{KT\}}(r) \cdot I_{\{i\}}(K_{|K|-r}). \end{aligned} \tag{2.6}$$

In view of (2.4)–(2.6), we obtain that

$$S_i(n) \cdot P_{i \to T} = \sum_{K \in \mathcal{C}} \sum_{\substack{r \in \{KT\} \\ 1 \leq r < |T|}} S_K(n + r) \cdot P_{T_r \to T_{r+1} \cdots T_{|T|}} \cdot I_{\{i\}}(K_{|K|-r}) + S_{i,T}(n + |T|).$$

Consequently,

$$\sum_{n=1}^{\infty} S_i(n) z^{-n} P_{i\to T} = \sum_{K\in\mathcal{C}} \sum_{\substack{r\in\{KT\}\\ 1\le r<|T|}} z^r \cdot P_{T_r\to T_{r+1}\cdots T_{|T|}} \cdot I_{\{i\}}(K_{|K|-r}) \cdot \sum_{n=1}^{\infty} S_K(n+r)\cdot z^{-n-r} + z^{|T|}\cdot \sum_{n=1}^{\infty} S_{i,T}(n+|T|)\cdot z^{-n-|T|}. \tag{2.7}$$

Note that for $r \in \{KT\}$ and $1 \le r < |T|$, we have $r < |K|$. So

$$\sum_{n=1}^{\infty} S_K(n+r)\cdot z^{-n-r} = f_K(z).$$

Hence we can rewrite (2.7) as

$$F_i(z)\cdot P_{i\to T} = \sum_{K\in\mathcal{C}} f_K(z) \cdot \sum_{\substack{r\in\{KT\}\\ 1\le r<|T|}} z^r \cdot P_{T_r\to T_{r+1}\cdots T_{|T|}} \cdot I_{\{i\}}(K_{|K|-r}) + z^{|T|}\cdot f_{i,T}(z). \tag{2.8}$$

Summing all $i \in \Delta$ gives

$$\sum_{i\in\Delta} F_i(z)\cdot P_{i\to T} = \sum_{K\in\mathcal{C}} f_K(z) \cdot \sum_{\substack{r\in\{KT\}\\ 1\le r<|T|}} z^r \cdot P_{T_r\to T_{r+1}\cdots T_{|T|}} + z^{|T|}\cdot \sum_{i\in\Delta} f_{i,T}(z). \tag{2.9}$$

Finally, combining (2.3) with (2.9), we conclude that

$$\sum_{i\in\Delta} F_i(z)\cdot P_{i\to T} = \sum_{K\in\mathcal{C}} f_K(z) \cdot \sum_{\substack{r\in\{KT\}\\ 1\le r<|T|}} z^r \cdot P_{T_r\to T_{r+1}\cdots T_{|T|}} + z^{|T|}\cdot f_T(z) - P_T.$$

Dividing by $P_{T_1\to T_2\cdots T_{|T|}}$ on both sides yields that

$$\sum_{i\in\Delta} F_i(z)\cdot P_{iT_1} = \sum_{K\in\mathcal{C}} f_K(z)\cdot \tilde{g}_{KT}(z) - \mu_{T_1}. \tag{2.10}$$

This, together with (2.2), completes the proof. □

Proposition 2.2. *The linear system* (2.1) *is nonsingular.*

Proof. W.l.o.g., suppose that $\Delta = \{1, \cdots, m\}$ and $\mathcal{C} = \{A, B, \cdots, T\}$. Let

$$Q(z) = \begin{pmatrix} P_{11}-z & P_{21} & \cdots & P_{m1} & -zX_A^{(1)} & -zX_B^{(1)} & \cdots & -zX_T^{(1)} \\ \cdots & & & & & & & \\ P_{1m} & P_{2m} & \cdots & P_{mm}-z & -zX_A^{(m)} & -zX_B^{(m)} & \cdots & -zX_T^{(m)} \\ P_{1A_1} & P_{2A_1} & \cdots & P_{mA_1} & -\tilde{g}_{AA}(z) & -\tilde{g}_{BA}(z) & \cdots & -\tilde{g}_{TA}(z) \\ \cdots & & & & & & & \\ P_{1T_1} & P_{2T_1} & \cdots & P_{mT_1} & -\tilde{g}_{AT}(z) & -\tilde{g}_{BT}(z) & \cdots & -\tilde{g}_{TT}(z) \end{pmatrix}.$$

Then we can rewrite (2.1) as

$$Q(z)\,(F_1(z), \cdots, F_m(z), f_A(z), \cdots, f_T(z))^T = (-\mu_1, \cdots, -\mu_m, -\mu_{A_1}, \cdots, -\mu_{T_1})^T.$$

Let $\varphi(z) = |Q(z)|$ be the determinant of $Q(z)$. It suffices to show that $\varphi(z)$ is a nonzero polynomial. Clearly, at the i-th row of $Q(z)$ with $1 \le i \le m$, the highest degree is 1 and occurs on the diagonal or after the m-th column; while at the j-th row with $j \ge m+1$, the highest degree polynomial occurs only on the diagonal. Therefore in the expansion of $\varphi(z)$, the unique highest degree monomial comes from the product of the diagonal terms. This, together with the fact the highest degree monomial of $\tilde{g}_{AA}(z)$ is $\frac{z^{|A|}}{P_{A_1\to A_2\cdots A_{|A|}}}$, implies that the unique highest degree monomial of $\varphi(z)$ is

$$(-1)^{m+|\mathcal{C}|}\frac{1}{P_{A_1\to A_2\cdots A_{|A|}}P_{B_1\to B_2\cdots B_{|B|}}\cdots P_{T_1\to T_2\cdots T_{|T|}}}z^{m+|A|+\cdots+|T|}.$$

It shows that $\varphi(z)$ is a nonzero polynomial as desired. □

For $i\in\Delta$ and $T\in\mathcal{C}$, let $F_i=F_i(1)$ and $f_T=f_T(1)$. Then $F_i=E\left(\sum\limits_{n<\tau}I_{\{Z_n=i\}}\right)$ is the mean staying time at i before τ, and $f_T=P(\tau=\tau_T<\infty)$ is the probability that the pattern T appears first among all the patterns in $\mathcal{C}$. Thus we have $E(\tau)=1+\sum\limits_{i\in\Delta}F_i$. Let $\tilde{g}_{KT}=\tilde{g}_{KT}(1)$. Substituting $z=1$ into Theorem 2.1 gives the following Corollary.

Corollary 2.3. *The following system of linear equations holds:*

$$\begin{cases}\sum\limits_{i\in\Delta}F_i\cdot P_{ij}=F_j+\sum\limits_{K\in\mathcal{C}}f_K\cdot X_K^{(j)}-\mu_j,\ j\in\Delta,\\ \sum\limits_{i\in\Delta}F_i\cdot P_{iT_1}=\sum\limits_{K\in\mathcal{C}}f_K\cdot\tilde{g}_{KT}-\mu_{T_1},\ T\in\mathcal{C}.\end{cases}\tag{2.11}$$

Remark 2.4. (1) For $z\geq 1$, define

$$F(z)=1+\sum_{i\in\Delta}F_i(z)=\sum_{n=0}^{\infty}P(\tau>n)\cdot z^{-n}$$

and

$$f(z)=\sum_{K\in\mathcal{C}}f_K(z)=\sum_{n=1}^{\infty}P(\tau=n)\cdot z^{-n}.$$

If we have solved all $f_K(z)$ with $K\in\mathcal{C}$, then we can obtain the generating function $f(z)$. In theory, we can obtain the distribution of τ. Particularly, we can calculate the moments of τ.

(2) Theorem 2.1 is the generalization of Theorem 3.3 of [7]. Summing all $j\in\Delta$ in the first part of (2.1), we get

$$(z-1)\cdot F(z)+z\cdot\sum_{K\in\mathcal{C}}f_K(z)=z.\tag{2.12}$$

In the case that $Z_1,Z_2,\cdots$ are i.i.d and $\mu_j>0$ for all j, $P_{ij}=\mu_j$ does not depend on i. Dividing by μ_{T_1} at the both side of the second part of (2.1) gives:

$$F(z)=\sum_{K\in\mathcal{C}}f_K(z)\cdot\tilde{g}_{KT}(z)/\mu_{T_1}.\tag{2.13}$$

If we define $c_{KT}(z)=\tilde{g}_{KT}(z)/(z\cdot\mu_{T_1})=\sum\limits_{r\in\{KT\}}\frac{z^{r-1}}{\mu_{T_1}\cdots\mu_{T_r}}$, then combining (2.12) with (2.13) yields Theorem 3.3 of [7]. Note that the definition of $c_{KT}(z)$ in [7] has a typo and we correct it here.

(3) To obtain $E(\tau)$ and $P(\tau=\tau_A)$ with $A\in\mathcal{C}$, we only need to solve one linear system involving $|\Delta|+|\mathcal{C}|$ equations and $|\Delta|+|\mathcal{C}|$ variables. Compared with the results in [4], [6] and [12], it is a much easy and effective way.

When $|T|=1$ and T is not a subpattern of K, we must have

$$\tilde{g}_{KT}(z)=\begin{cases}0,\ K\neq T,\\ z,\ K=T.\end{cases}$$

If $j \in \mathcal{C}$, then $F_j(z) = 0$. By the above discussion, Theorem 2.1 yields the following Corollary.

Corollary 2.5. *If the lengths of all patterns in $\mathcal{C}$ are 1, then the following linear system holds:*

$$\begin{cases} \sum\limits_{i \notin \mathcal{C}} F_i(z) \cdot P_{ij} = z \cdot f_j(z) - \mu_j, \ j \in \mathcal{C}, \\ \sum\limits_{i \notin \mathcal{C}} F_i(z) \cdot P_{ij} = z \cdot F_j(z) - \mu_j, \ j \notin \mathcal{C}. \end{cases}$$

When all pattern contains only one element, we only need to solve a linear system involving $|\Delta|$ equations.

Corollary 2.6. *Suppose that the first elements of all patterns in $\mathcal{C}$ are equal and A is any pattern in $\mathcal{C}$. Then the following linear system holds:*

$$\begin{cases} \sum\limits_{K \in \mathcal{C}} f_K = 1, \\ \sum\limits_{K \in \mathcal{C}} f_K \cdot (\tilde{g}_{KT} - \tilde{g}_{KA}) = 0, \ T \in \mathcal{C}, \ T \neq A. \end{cases} \tag{2.14}$$

Proof. Set $h = A_1$. Then $T_1 = h$ for all $T \in \mathcal{C}$. In this case, the second part of (2.11) can be rewritten as following:

$$\sum_{i \in \Delta} F_i \cdot P_{ih} = \sum_{K \in \mathcal{C}} f_K \cdot \tilde{g}_{KT} - \mu_h, \ T \in \mathcal{C}.$$

It shows that for all $T \in \mathcal{C}$, the values $\sum\limits_{K \in \mathcal{C}} f_K \cdot \tilde{g}_{KT}$ are the same. Particularly,

$$\sum_{K \in \mathcal{C}} f_K \cdot \tilde{g}_{KT} = \sum_{K \in \mathcal{C}} f_K \cdot \tilde{g}_{KA}.$$

This, combining with the fact that $\sum_{K \in \mathcal{C}} f_K = 1$ yields our result. □

When the first elements of all patterns are equal, namely h, the calculation become more simplified. To solve f_K with $K \in \mathcal{C}$, it is enough to solve a linear system of $|\mathcal{C}|$ equations. In this case, the stopping probabilities are only related to the transition probability among those states in Δ_1, but neither the initial distribution nor the transition probability P_{ij} with i or j outside Δ_1, where Δ_1 is the set of elements of patterns in $\mathcal{C}$. This is actually true. Intuitively, all patterns do not occur before the first visiting h. In addition, if the process stays outside Δ_1 and no pattern has occurred, then the behavior before his next visiting h will not affect the stopping probabilities.

Sometimes we are interested in when the distribution of Z_τ is the same as the initial distribution. The Corollary below gives the answer.

Corollary 2.7. *Assume that $\{Z_n\}$ is irreducible and has the unique stationary distribution π.*

(1) *The distribution of Z_τ is the same as the initial distribution if and only if there is a constant c such that $F_i = c \cdot \pi_i$ for all $i \in \Delta$. Actually, $c = (\sum_{K \in \mathcal{C}} f_K \cdot \tilde{g}_{KT} - \mu_{T_1})/\pi_{T_1}$ with any given $T \in \mathcal{C}$, and $E(\tau) = 1 + c$.*

(2) *If the distribution of Z_τ is the same as the initial distribution, then the following linear system holds:*

$$\begin{cases} \sum\limits_{K \in \mathcal{C}} f_K = 1, \\ \sum\limits_{K \in \mathcal{C}} f_K \cdot (\tilde{g}_{KT} - X_K^{(T_1)}) = c \cdot \pi_{T_1}, \ T \in \mathcal{C}. \end{cases} \tag{2.15}$$

Proof. By (1) and Corollary 2.3, (2) follows immediately. Thus we only need to prove (1). The first part of (2.11) shows that the distribution of Z_τ is the same as the initial distribution if and only if

$$\sum_{i\in\Delta} F_i \cdot P_{ij} = F_j,\ j \in \Delta. \tag{2.16}$$

Equivalently, there is a constant c such that $F_i = c\cdot\pi_i$ for all $i\in\Delta$. In this case, $E(\tau) = 1+\sum_{i\in\Delta} F_i = 1+c$. By (2.16) and the second part of (2.11), we have

$$F_{T_1} = \sum_{K\in\mathcal{C}} f_K \cdot \tilde{g}_{KT} - \mu_{T_1}.$$

It follows that $c = (\sum_{K\in\mathcal{C}} f_K\cdot\tilde{g}_{KT} - \mu_{T_1})/\pi_{T_1}$ as desired. □

§3 Examples

We begin with the analysis of Example 1 of [12]. The mean waiting time and the generating function of τ are calculated in Example 1 and Example 3 of [12] respectively, while the stopping probability is obtained in Example 3.1 of [4]. We now recalculate all these values by applying our results.

Example 3.1. Suppose that $\Delta = \{1,2,3\}, \mathcal{C} = \{323, 313, 33\}, \mu_1 = \mu_2 = \mu_3 = 1/3$ and the one-step transition probability matrix is

$$P = \begin{pmatrix} 3/4 & 0 & 1/4 \\ 0 & 3/4 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}.$$

Let $A = 323$,$B = 313$ and $C = 33$. By calculation, we get

$$\tilde{g}_{AA}(z) = z+16z^3,\quad \tilde{g}_{BA} = z,\quad \tilde{g}_{CA} = z,$$
$$\tilde{g}_{AB}(z) = z,\quad \tilde{g}_{BB}(z) = z+16z^3,\quad \tilde{g}_{CB} = z,$$
$$\tilde{g}_{AC} = z,\quad \tilde{g}_{BC} = z,\quad \tilde{g}_{CC} = z+2z^2.$$

Put these values into (2.1), we get

$$\begin{pmatrix} \frac{3}{4}-z & 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{3}{4}-z & \frac{1}{4} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2}-z & -z & -z & -z \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & -z-16z^3 & -z & -z \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & -z & -z-16z^3 & -z \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & -z & -z & -z-2z^2 \end{pmatrix} \begin{pmatrix} F_1(z) \\ F_2(z) \\ F_3(z) \\ f_A(z) \\ f_B(z) \\ f_C(z) \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \end{pmatrix}.$$

It is easily seen that

$$f_A(z) = f_B(z) = \frac{F_3(z)}{16z^2}, f_C(z) = \frac{F_3(z)}{2z}, \text{ and } F_1(z) = F_2(z) = \frac{4+3F_3(z)}{12z-9}.$$

In addition, $F_3(z) = 8z(4z-1)/(96z^3-72z^2-9)$. Therefore

$$E(z^{-\tau}) = f(z) = f_A(z)+f_B(z)+f_C(z) = \frac{16z^2-1}{3z(32z^3-24z^2-3)}.$$

Writing $z = 1/\alpha$ yields that $E(\alpha^\tau) = \frac{\alpha^2(\alpha^2-16)}{3(3\alpha^3+24\alpha-32)}$. Taking $z=1$ gives $f_A = f_B = 1/10, f_C =$

$8/10, F_1 = F_2 = 44/15, F_3 = 24/15$, and hence $E(\tau) = 1 + F_1 + F_2 + F_3 = 127/15$. These results are all in agreement with that in [4] and [12].

Another way is to apply Corollary 2.6 and Corollary 2.7. Because the first elements of A, B, C are equal, substituting

$$\tilde{g}_{AA} = 17,\ \tilde{g}_{BA} = 1,\ \tilde{g}_{CA} = 1$$
$$\tilde{g}_{AB} = 1,\ \tilde{g}_{BB} = 17,\ \tilde{g}_{CB} = 1$$
$$\tilde{g}_{AC} = 1,\ \tilde{g}_{BC} = 1,\ \tilde{g}_{CC} = 3$$

into (2.14) yields the following linear system:

$$\begin{cases} f_A + f_B + f_C = 1, \\ -16 \cdot f_A + 16 \cdot f_B = 0, \\ -16 \cdot f_A + 2 \cdot f_C = 0. \end{cases}$$

Thus $f_A = f_B = 1/10$ and $f_C = 8/10$. It is easy to see that the stationary distribution is $\pi_1 = \pi_2 = \pi_3 = 1/3$. Because the last elements of A, B, C are all equal to 3, by Corollary 2.7,

$$E(\tau|Z_1 = 3) = 1 + (f_A \cdot \tilde{g}_{AA} + f_B \cdot \tilde{g}_{BA} + f_C \cdot \tilde{g}_{CA} - 1)/\pi_3 = 29/5.$$

Clearly, $P(\tau_3 = 1) = \frac{1}{3}$ and $P(\tau_3 = n) = \frac{2}{3} \cdot (\frac{3}{4})^{n-2} \cdot \frac{1}{4}$ for $n \geq 2$. Therefore

$$E(\tau) = E(\tau_3) - 1 + E(\tau|Z_1 = 3) = 127/15.$$

Example 3.2. Suppose that $\Delta = \{1, 2\}, \mathcal{C} = \{A, B\}, A = 22, B = 121$ and

$$P = \begin{pmatrix} 1/4 & 3/4 \\ 3/4 & 1/4 \end{pmatrix}.$$

When will the distribution of Z_τ be the same as the initial distribution?

By calculating, we get $\tilde{g}_{AA} = 5$, $\tilde{g}_{BA} = 0$, $\tilde{g}_{AB} = 0$ and $\tilde{g}_{BB} = 25/9$. The stationary distribution is $\pi_1 = \pi_2 = 1/2$. Using Corollary 2.7, we have

$$\begin{cases} f_A + f_B = 1, \\ 4 \cdot f_A = \frac{1}{2} \cdot c, \\ \frac{16}{9} \cdot f_B = \frac{1}{2} \cdot c. \end{cases}$$

Hence $\mu_2 = f_A = 4/13, \mu_1 = f_B = 9/13$ and $c = 32/13$. In addition, $F_1 = c \cdot \pi_1 = 16/13$, $F_2 = c \cdot \pi_2 = 16/13$ and $E(\tau) = 1 + c = 45/13$.

Acknowledgement. The authors wish to express their thanks to the referee for his/her constructive suggestions and many helpful comments.

References

[1] J Brofos. *A Markov chain analysis of a pattern matching coin game*, arXiv preprint, arXiv:1406.2212, 2014.

[2] O Chrysaphinou, S Papastavridis. *The occurrence of sequence patterns in repeated dependent experiments*, Theory Probab Appl, 1991, 35(1): 145-152.

[3] J C Fu, Y M Chang. *On probability generating functions for waiting time distributions of compound patterns in a sequence of multistate trials*, J Appl Probab, 2002, 39: 70-80.

[4] R J Gava, D Salotti. *Stopping probabilities for patterns in Markov chains*, J Appl Probab, 2014, 51(1): 287-292.

[5] H U Gerber, S R Li. *The occurrence of sequence patterns in repeated experiments and hitting times in a Markov chain*, Stochastic Process Appl, 1981, 11(1): 101-108.

[6] J Glaz, M Kulldorff, V Pozdnyakov, J M Steele. *Gambling teams and waiting times for patterns in two-state Markov chains*, J Appl Probab, 2006, 43(1): 127-140.

[7] L J Guibas, A M Odlyzko. *String overlaps, pattern matching, and nontransitive games*, J Combin Theory Ser A, 1981, 30(2): 183-208.

[8] S R Li. *A martingale approach to the study of occurrence of sequence patterns in repeated experiments*, Ann Probab, 1980, 8(6): 1171-1176.

[9] J I Naus. *The distribution of the size of the maximum cluster of points on a line*, J Amer Statist Assoc, 1965, 60(310): 532-538.

[10] J I Naus, V T Stefanov. *Double-scan statistics*, Methodol Comput Appl Probab, 2002, 4(2): 163-180.

[11] Y Nishiyama. *Pattern matching probabilities and paradoxes as a new variation on Penney's coin game*, Int J Pure Appl Math, 2010, 59(3): 357-366.

[12] V Pozdnyakov. *On occurrence of patterns in Markov chains: method of gambling teams*, Statist Probab Lett, 2008, 78(16): 2762-2767.

[1] School of Mathematical Sciences, Zhejiang University, Hangzhou 310027, China.
Email: zhaomz@zju.edu.cn, xudong_1236@163.com

[2] Department of Mathematics, Hangzhou Normal University, Hangzhou 310036, China.
Email: zhanghz789@163.com