On Occurrences of Words under Markovian Hypothesis

Ourania Chryssaphinou[1], Margarita Karaliopoulou[1], and Nikolaos Limnios[2]

[1] University of Athens
Department of Mathematics,
157 84, Athens,
Greece
(e-mail: ocrysaf@cc.uoa.gr, mkaraliop@math.uoa.gr)

[2] Universite de Technologie de Compiegne
Laboratoire de Mathematiques Appliquees Centre de Recherches de Royallieu
BP 20529,
60205 COMPIEGNE CEDEX, France
(e-mail: Nikolaos.Limnios@utc.fr)

Abstract. We consider a finite set of words $W = \{w_1, w_2, \ldots, w_\nu\}$ which are produced under the Markovian hypothesis. We study the distances between word occurrences and we give explicit formulae for the corresponding distributions in the case of having words of equal lengths.The obtained results can be applied to certain problems concerning DNA sequences, as well as, general sequential analysis.

Keywords: Word, Markov chain, distance between occurrences, semi-Markov, waiting time.

1 Preliminaries

Consider an alphabet $\Omega = \{\alpha_1, \ldots, \alpha_\ell\}$ with $\ell \geq 2$. We call word a finite sequence of elements of Ω. Let $W = \{w_1, \ldots, w_\nu\}$ a finite sets of words where $w_i = (\alpha_{i_1}, \ldots, \alpha_{i_{k_i}})$, $\alpha_{i_{n_i}} \in \Omega$, $n_i = 1, \ldots, k_i$ where k_i denotes the length of word w_i and let $k_i > 1$. We assume that the set of words is reduced. Let us consider a sequence of outcomes $\{J_n^*\}_{n\geq 1}$ generated by a Markov chain with state space Ω, and let $\mathbf{P} = (\wp(\alpha_i, \alpha_j))_{\alpha_i, \alpha_j \in \Omega}$, the transition probability matrix. We write $\mathbf{P}_l^n = (\wp^n(\alpha_i, \alpha_j))_{\alpha_i, \alpha_j \in \Omega}$, where

$$\wp^n(\alpha_i, \alpha_j) = P(J_{n+1}^* = \alpha_j | J_1^* = \alpha_i).$$

A word w_i occurs at time γ iff $J_{\gamma-k_i+1}^* = \alpha_{i_1}, \ldots, J_\gamma^* = \alpha_{i_{k_i}}$.

Definition 1 *Let W_λ a subset of W. We define*

$$U^* = \min\{\gamma \geq 1 : \ \textit{a word occurs at}\ \gamma\},$$

$$M_{W_\lambda}^* = \min\{\gamma \geq 1 : \ \textit{a word from the subset}\ W_\lambda\ \textit{occurs at}\ \gamma\},$$

and let y_0 be the first word which appears.

Clearly, the variable U^* indicates the waiting time (number of letters) for the first occurrence of any word from the set W, while the variable $M^*_{W_\lambda}$ indicates the waiting time (number of letters) for the first occurrence of any word from the subset W_λ. In section 2 we assume words of the same length and we obtain explicit formulae concerning the distributions of the above random variables. In section 3, under the same assumption we model the process of word occurrences via a semi-Markov model for which we derive the kernel as well as relative results. The under consideration random variables are of great interest for the study of biological sequences where the corresponding alphabet is $\Omega = \{A, C, G, T\}$.

Recurrent relations for variable $M^*_{\{w_i\}}$ are given by Blom and Thorburn (1982) in the i.i.d case, by Chryssaphinou and Papastavridis (1990) and Robin and Daudin (1999) in the Markov case.

2 Words of the same length and without a word at the beginning of the sequence

Let us examine the case where $k_i = k$, $\forall\, w_i \in W$. We construct a new Markov Chain $\{X_n^*\}_{n\geq 1}$ where

$$X_n^* = (J_n^*, \ldots, J_{n+k-1}^*),\ n \geq 1, \tag{1}$$

with state space $\quad \Omega^k = \Omega \times \ldots \times \Omega \quad$ and

$$u_i = (\alpha_1^{u_i}, \ldots, \alpha_k^{u_i}),\ \forall\, i = 1, \ldots, \ell^k \quad \text{and} \quad \alpha_n^{u_i} \in \Omega, \quad \forall\, n = 1, \ldots, k,$$

The new transition matrix is

$$\tilde{P} = (\tilde{p}(u_i, u_j)),\ \ u_i,\ u_j \in \Omega^k, \tag{2}$$

where

$$\begin{aligned}
\tilde{p}(u_i, u_j) &= \mathbb{P}(X_{n+1}^* = u_j | X_n^* = u_i) \\
&= \mathbb{P}(J_{n+1}^* = \alpha_1^{u_j}, \ldots, J_{n+k}^* = \alpha_k^{u_j} \mid J_n^* = \alpha_1^{u_i}, \ldots, J_{n+k-1}^* = \alpha_k^{u_i}) \\
&= I_{\{\alpha_2^{u_i} = \alpha_1^{u_j}, \ldots, \alpha_k^{u_i} = \alpha_{k-1}^{u_j}\}} \wp(\alpha_{k-1}^{u_j}, \alpha_k^{u_j}).
\end{aligned} \tag{3}$$

The initial distribution is

$$\begin{aligned}
\mathbb{P}(X_1^* = u_i) &= \mathbb{P}(J_1^* = \alpha_1^{u_i}, \ldots, J_k^* = \alpha_k^{u_i}) \\
&= \sigma(\alpha_1^{u_i}) \wp(\alpha_1^{u_i}, \alpha_2^{u_j}), \ldots, \wp(\alpha_{k-1}^{u_i}, \alpha_k^{u_i}),
\end{aligned} \tag{4}$$

where σ is the initial distribution of Markov chain J^*. We note

$$\tilde{P}_1 = (\mathbb{P}(X_1^* = u_1), \ldots, \mathbb{P}(X_1^* = u_{\ell^k})). \tag{5}$$

Since $W \subseteq \Omega^k, \exists\, r_1, \ldots, r_\nu \in \{1, \ldots, \ell^k\} : w_1 = u_{r_1}, \ldots, w_\nu = u_{r_\nu}$. Let $B^c = \Omega^k \setminus B, \quad \forall\, B \subseteq \Omega^k$. We define the matrices

$$\tilde{P}_{B^c B^c}, \quad \tilde{P}_{B^c B}, \tag{6}$$

which are the restriction of the transition matrices $\tilde{P}$ in $B^c \times B^c$ and $B^c \times B$ respectively. Generally $\forall\, B_1, B_2 \subseteq \Omega^k$, let $\tilde{P}_{B_1 B_2}$ the restriction of $\tilde{P}$ in $B_1 \times B_2$.

For $W_\lambda \subseteq W$, where $\mid W_\lambda \mid = \lambda$, we now define the nth-order transition matrix

$$\tilde{P}^n_{W^c_\lambda W^c_\lambda} = (\tilde{p}^n_{W^c_\lambda W^c_\lambda}(u_i, u_j)), \;\; u_i,\; u_j \in W^c_\lambda, \tag{7}$$

where $\tilde{p}^n_{W^c_\lambda W^c_\lambda}(u_i, u_j) = \mathbb{P}(X^*_{n+1} = u_j, X^*_n \in W^c_\lambda, \ldots, X^*_2 \in W^c_\lambda | X^*_1 = u_i)$, and $\mathbb{P}^1_{W^c_\lambda W^c_\lambda} = \mathbb{P}_{W^c_\lambda W^c_\lambda}, \quad \mathbb{P}^0_{W^c_\lambda W^c_\lambda} = \mathbb{I}_{\ell^k - \lambda}$. Finally, let us define

$$\tilde{P}_{W_\lambda} = (\mathbb{P}(X^*_1 = u_i)),\; u_i \in W_\lambda, \qquad \tilde{P}_{W^c_\lambda} = (\mathbb{P}(X^*_1 = u_i)), u_i \notin W_\lambda. \tag{8}$$

Now we are ready to present the following results.

Proposition 1 *With the above notation the distribution of the random variable $M^*_{W_\lambda}$ is given by*

$$\mathbb{P}(M^*_{W_\lambda} = n) = \begin{cases} 0, & n < k, \\ \tilde{P}_{W_\lambda} \boldsymbol{1}'_\lambda, & n = k, \\ [\tilde{P}_{W^c_\lambda}] \times [\tilde{P}^{n-k-1}_{W^c_\lambda W^c_\lambda}] \times [\tilde{P}_{W^c_\lambda W_\lambda}] \boldsymbol{1}'_\lambda, & n > k. \end{cases} \tag{9}$$

where $\boldsymbol{1}_\lambda = (1, \ldots, 1)$, $(1 \times \lambda$ *matrix*$)$

Proof. It is $\mathbb{P}(M^*_{W_\lambda} = k) = \mathbb{P}(X^*_1 \in W_\lambda) = \tilde{P}_{W_\lambda} \mathbf{1}'_\lambda$. For $n > k$ we have

$$\begin{aligned}
&\mathbb{P}(M^*_{W_\lambda} = n) = \mathbb{P}(X^*_{n-k+1} \in W_\lambda, X^*_{n-k} \in W^c_\lambda, \ldots, X^*_2 \in W^c_\lambda, X^*_1 \in W^c_\lambda) \\
&= \sum_{u_i \in W^c_\lambda} \mathbb{P}(X^*_{n-k+1} \in W_\lambda, X^*_{n-k} \in W^c_\lambda, \ldots, X^*_2 \in W^c_\lambda \mid X^*_1 = u_i) \\
&\qquad \mathbb{P}(X^*_1 = u_i) \\
&= \sum_{u_i \in W^c_\lambda} \sum_{u_j \in W_\lambda} \mathbb{P}(X^*_{n-k+1} = u_j \mid X^*_{n-k} \in W^c_\lambda, \ldots, X^*_2 \in W^c_\lambda, X^*_1 = u_i) \\
&\qquad \mathbb{P}(X^*_{n-k} \in W^c_\lambda, X^*_{n-k-1} \in W^c_\lambda \ldots, X^*_2 \in W^c_\lambda \mid X^*_1 = u_i) \mathbb{P}(X^*_1 = u_i) \\
&= \sum_{u_i \in W^c_\lambda} \sum_{u_j \in W_\lambda} \sum_{u_l \notin W_\lambda} \mathbb{P}(X^*_{n-k+1} = u_j \mid X^*_{n-k} = u_l) \\
&\qquad \mathbb{P}(X^*_{n-k} = u_l, X^*_{n-k-1} \in W^c_\lambda \ldots, X^*_2 \in W^c_\lambda \mid X^*_1 = u_i) \mathbb{P}(X^*_1 = u_i) \\
&= [\tilde{P}_{W^c_\lambda}] \times [\tilde{P}^{n-k-1}_{W^c_\lambda W^c_\lambda}] \times [\tilde{P}_{W^c_\lambda W_\lambda}] \mathbf{1}'_\lambda,
\end{aligned}$$

which completes the proof.

Proposition 2 *For every $w_i \in W$ the following is valid*

$$\mathbb{P}(U^* = \gamma,\, y_0 = w_i) = \begin{cases} \tilde{P}_1 \; e'_{\ell^k;r_i}, & \gamma = k, \\ \tilde{P}_{W^c}\,[\tilde{P}_{W^cW^c}]^{\gamma-k-1}\,\tilde{P}_{W^c\Omega^k}\, e'_{\ell^k;r_i}, & \gamma > k, \end{cases} \tag{10}$$

where $e_{n;m} = (0, \ldots, \underbrace{1}_{m-position}, \ldots, 0)$ *($1 \times n$ matrix)*

Proof. The case of $\gamma = k$ is obvious, since $\mathbb{P}(U^* = k, y_0 = w_i) = \mathbb{P}(X_1^* = u_{r_i})$. For $\gamma > k$ we proceed as follows

$$\begin{aligned}
&\mathbb{P}(U^* = \gamma, y_0 = w_i) \\
&= \mathbb{P}(X^*_{\gamma-k+1} = u_{r_i}, X^*_{\gamma-k} \in W^c, \ldots, X_2^* \in W^c, X_1^* \in W^c) \\
&= \sum_{u_s \notin W_\lambda} \mathbb{P}(X^*_{\gamma-k+1} = u_{r_i}, X^*_{\gamma-k} \in W^c, \ldots, X_2^* \in W^c \mid X_1^* = u_s) \\
&\quad\; \mathbb{P}(X_1^* = u_s) \\
&= \sum_{u_s \notin W} \mathbb{P}(X^*_{\gamma-k+1} = u_{r_i} \mid X^*_{\gamma-k} \in W^c, \ldots, X_2^* \in W^c, X_1^* = u_s) \\
&\quad\; \mathbb{P}(X^*_{\gamma-k} \in W^c, X^*_{\gamma-k-1} \in W^c \ldots, X_2^* \in W^c \mid X_1^* = u_s) \\
&\quad\; \mathbb{P}(X_1^* = u_s) \\
&= \sum_{u_s \notin W} \sum_{u_l \notin W} \mathbb{P}(X^*_{\gamma-k+1} = u_{r_i} \mid X^*_{\gamma-k} = u_l) \\
&\quad\; \mathbb{P}(X^*_{\gamma-k} = u_l, X^*_{\gamma-k-1} \in W^c \ldots, X_2^* \in W^c \mid X_1^* = u_s) \\
&\quad\; \mathbb{P}(X_1^* = u_s) \\
&= \tilde{P}_{W^c}\; [\tilde{P}_{W^cW^c}]^{\gamma-k-1}\; \tilde{P}_{W^c\Omega^k}\; e'_{q^k;r_i},
\end{aligned}$$

which ends the proof.

3 Words of the same length and with a word at the beginning of the sequence

We now consider $\{J_n\}$ where $J_n = J^*_{U^*+n}$, $\forall\, n \geq -U^* + 1$. We want to study the sequence $J_0, J_1, \ldots$ under the assumption that the word w_i has occurred with probability $\theta_i = \mathbb{P}(J_{-k_i+1} = \alpha_{i_1}, \ldots, J_0 = \alpha_{i_{k_i}})$, $i = 1, \ldots, \nu$. Without loss of generality we can take $\theta_i = \mathbb{P}(y_0 = w_i)$, $i = 1, \ldots, \nu$.

The sequence $\{J_n, n \geq 0\}$ is a Markov chain with first order transition probabilities $\mathbb{P}(J_{n+1} = \alpha_j | J_n = \alpha_i) = \wp(\alpha_i, \alpha_j)$ and $\mathbb{P}(J_0 = \alpha_\zeta) = \sum_{i=1}^{\nu} I_{\{\alpha_{i_{k_i}} = \alpha_\zeta\}} \theta_i$.

In this case a word w_i occurs at time γ iff $J_{\gamma-k_i+1} = \alpha_{i_1}, \ldots, J_\gamma = \alpha_{i_{k_i}}$.

3.1 The Semi-Markov Model

In the case where words do not overlap, Biggins and Cannings (1987), introduced the idea of modelling the process of word occurrences via a semi-Markov model. Recently, Robin and Daudin (2001) generalized the idea considering the fact that words may overlap. Our aim is to determine the kernel of this new process. In order to present our results we need some more definitions and notations.

Definition 2 *Let us define the stochastic processes* $\{U_n,\ n \geq 0\}, \{y_n,\ n \geq 0\}$*, which describe the times of word occurrences and the corresponding words respectively, where* $U_0 = 0$ *and*

$$U_n = \min\{\gamma > U_{n-1} : a\ word\ from\, W\ occurs\ at\, \gamma\}, \quad n \geq 1, \tag{11}$$

$$y_n = w_i, \quad w_i \in W,\ i = 1, \ldots, \ell. \tag{12}$$

The process $\{(y_n, U_n), n \in \mathbb{N}\}$ is an homogenous Discrete time Markov Renewal Process(DTMRP) since

$$\begin{aligned}
&\mathbb{P}(y_{n+1} = w_j, U_{n+1} - U_n = \gamma \,|\, y_0, \ldots, y_n = w_i, U_0, \ldots, U_n) = \\
&\mathbb{P}(y_{n+1} = w_j, U_{n+1} - U_n = \gamma \,|\, y_n = w_i) = \\
&\mathbb{P}(y_1 = w_j, U_1 = \gamma \,|\, y_0 = w_i) \ = \ q_{ij}(\gamma), \quad \forall n \geq 1
\end{aligned}$$

Let us consider the following notation

- $\mathcal{M}_E$, the set of non negative matrices on $E \times E$.
- $\mathbf{I}_E \in \mathcal{M}_E$, the identity matrix, $\underset{A}{0_\circ} \in \mathcal{M}_E$, the null matrix.
- $\mathcal{M}_E(\mathbb{N})$, the set of matrix-valued functions: $\mathbb{N} \to \mathcal{M}_E$. If $A \in \mathcal{M}_E(\mathbb{N})$,we have $A = (A(\gamma) : \gamma \in \mathbb{N})$, where for fixed $\gamma \in \mathbb{N}, A(\gamma) = (A_{ij}(\gamma) : i, j \in E) \in \mathcal{M}_E$.

Then $q \in \mathcal{M}_E(\mathbb{N})$ $(E = \{1, \ldots \nu\})$ is the discrete time semi-Markov kernel relevant to the DTMRP $\{(y_n, U_n), n \in \mathbb{N}\}$. We have

$$q_{ij}^{(r)}(\gamma) = \mathbb{P}(y_r = w_j,\ U_r = \gamma \,|\, y_0 = w_i), \tag{13}$$

where $q^{(r)}$ is the r- fold convolution of q .Then we can define

$$\psi_{ij}(\gamma) = \sum_{r=0}^{\gamma} q_{ij}^{(r)}(\gamma), \quad w_i, w_j \in W,\ \gamma \in \mathbb{N}. \tag{14}$$

We can write

$$\psi_{ij}(\gamma) = q_{ij}(\gamma) + \sum_{s=1}^{\nu} \sum_{z=1}^{\gamma-1} \psi_{is}(z) q_{sj}(\gamma - z), \quad \text{for } \gamma \geq 1. \tag{15}$$

Definition 3 *For all* $r \in \mathbb{N}^*, \forall w_i, w_j \in W$ *let* $M_{ij}^{(r)}$ *be the number of letters of the r-th occurrence of* w_j *after* w_i*'s occurrence.*

Definition 4 *For all* $W_\lambda \subseteq W$ *we define*

$$M_{iW_\lambda} = \min_{n\geq 1}\{U_n : y_n \in W_\lambda\} \text{ under the event } \{y_0 = w_i\}. \tag{16}$$

We will note M_{ij} *for* $M_{i\{w_j\}}$. *Obviously* $M_{ij} = M_{ij}^{(1)}$.

If $g_{ij}(\gamma) = \mathbb{P}(M_{ij} = \gamma)$ and $g_{ij}^{(r)}(\gamma) = \mathbb{P}(M_{ij}^{(r)} = \gamma)$, then

$$\psi_{ij}(\gamma) = \begin{cases} \sum_{r=0}^{\gamma} g_{jj}^{(r)}(\gamma), & i = j \\ \sum_{r=0}^{\gamma} g_{ij} * g_{jj}^{(r)}(\gamma), & i \neq j. \end{cases} \tag{17}$$

Definition 5 *We assume* $k_i = k$ *for all* $w_i \in W$. *We define*

$$X_n = (J_{n-k+1}, \ldots, J_n), \quad n \geq 0, \tag{18}$$

with

$$\mathbb{P}(X_0 = u_{r_i}) = \mathbb{P}(y_0 = w_i) = \mathbb{P}(J_{-k+1} = \alpha_{i_1}, \ldots, J_0 = \alpha_{i_k}),$$
$$\mathbb{P}(X_0 = u_i) = 0, \; \forall\, u_i \notin W.$$

Clearly, $\{X_n, n \geq 0\}$ is a Markov Chain with state space $\Omega^k = \{u_1, \ldots, u_{\ell^k}\}$, where $\forall\; i = 1, \ldots, \ell^k$, $u_i = (\alpha_1^{u_i}, \ldots, \alpha_k^{u_i})$, $\alpha_\zeta^{u_i} \in \Omega$, $\forall\, \zeta = 1, \ldots, k$ and $\tilde{P} = (\tilde{p}(u_i, u_j)) = (\mathbb{P}(X_{n+1} = u_j | X_n = u_i)), \forall\, u_i, u_j \in \Omega^k$.

Using the above definitions and notations we obtain the following results:

Proposition 3 *For every* $w_i, w_j \in W$ *we have:*

$$q_{ij}(\gamma) = \begin{cases} e_{\ell^k; r_i}\, \tilde{P}\, e'_{\ell^k; r_j}, & \gamma = 1 \\ e_{\ell^k; r_i}[\tilde{P}_{\Omega^k W^c}][\tilde{P}_{W^c W^c}]^{\gamma-2}[\tilde{P}_{W^c \Omega^k}]e'_{\ell^k; r_j}, & \gamma \geq 2 \end{cases} \tag{19}$$

Proof. It is

$$q_{ij}(1) = \mathbb{P}(X_1 = w_j | X_0 = w_i) = e_{\ell^k; r_i}\, \tilde{P}\, e'_{\ell^k; r_j}.$$

For $\gamma \geq 2$ we have

$$q_{ij}(\gamma) = \mathbb{P}(X_\gamma = w_j, X_1 \notin W, \ldots, X_{\gamma-1} \notin W \mid X_0 = w_i)$$
$$= \mathbb{P}(X_\gamma = u_{r_j}, X_1 \notin W, \ldots, X_{\gamma-1} \notin W \mid X_0 = u_{r_i})$$
$$= \sum_{u_n, u_s \notin W}$$
$$\mathbb{P}(X_\gamma = u_{r_j}, X_{\gamma-1} = u_n, X_{\gamma-2} \notin W, \ldots, X_2 \notin W, X_1 = u_s | X_0 = u_{r_i})$$
$$= \sum_{u_n, u_s \notin W}$$
$$\mathbb{P}(X_\gamma = u_{r_j}, X_{\gamma-1} = u_n, X_{\gamma-2} \notin W, \ldots, X_2 \notin W, | X_1 = u_s, X_0 = u_{r_i})$$
$$\mathbb{P}(X_1 = u_s | X_0 = u_{r_i})$$

that is

$$
\begin{aligned}
&q_{ij}(\gamma)\\
&= \sum_{u_n,u_s\notin W} \mathbb{P}(X_\gamma = u_{r_j}, X_{\gamma-1} = u_n, X_{\gamma-2} \notin W, \ldots, X_2 \notin W, |X_1 = u_s)\\
&\quad \mathbb{P}(X_1 = u_s | X_0 = u_{r_i})\\
&= \sum_{u_n,u_s\notin W} \mathbb{P}(X_\gamma = u_{r_j} | X_{\gamma-1} = u_n, X_{\gamma-2} \notin W, \ldots, X_2 \notin W, X_1 = u_s)\\
&\quad \mathbb{P}(X_{\gamma-1} = u_n, X_{\gamma-2} \notin W, \ldots, X_2 \notin W, |X_1 = u_s)\mathbb{P}(X_1 = u_s | X_0 = u_{r_i})\\
&= \sum_{u_n,u_s\notin W} \mathbb{P}(X_\gamma = u_{r_j} | X_{\gamma-1} = u_n)\\
&\mathbb{P}(X_{\gamma-1} = u_n, X_{\gamma-2} \notin W, \ldots, X_2 \notin W, |X_1 = u_s)\mathbb{P}(X_1 = u_s | X_0 = u_{r_i})\\
&= e_{\ell^k;r_i}[\tilde{P}_{\Omega^k W^c}][\tilde{P}_{W^c W^c}]^{\gamma-2}[\tilde{P}_{W^c \Omega^k}]e'_{\ell^k;r_j}\ .
\end{aligned}
$$

Proposition 4 *It is*

$$
\mathbb{P}(M_{iW_\lambda} = \gamma) = \begin{cases} e_{\ell^k;r_i}\ \tilde{P}_{\Omega^k W_\lambda}\ \mathbf{1}'_\lambda{}', & \gamma = 1\\ e_{\ell^k;r_i}\ \tilde{P}_{\Omega^k W_\lambda^c}[\tilde{P}_{W_\lambda^c W_\lambda^c}]^{\gamma-2}\tilde{P}_{W_\lambda^c W_\lambda}\,\mathbf{1}'_\lambda, & \gamma \geq 2. \end{cases} \tag{20}
$$

Proof. The random variable M_{iW_λ} can be expressed as follows

$$
M_{iW_\lambda} = \min\{n \geq 1 : X_n \in W_\lambda\} \text{ over } \{X_0 = w_i\}. \tag{21}
$$

If $\gamma = 1$, then

$$
\mathbb{P}(M_{iW_\lambda} = 1) = \mathbb{P}(X_1 \in W_\lambda \mid X_0 = w_i) = e_{\ell^k;r_i}\ \tilde{P}_{\Omega^k W_\lambda}\ \mathbf{1}'_\lambda.
$$

If $\gamma \geq 2$, then

$$
\begin{aligned}
&\mathbb{P}(M_{iW_\lambda} = \gamma) = \mathbb{P}(X_\gamma \in W_\lambda, X_{\gamma-1} \notin W_\lambda, \ldots, X_1 \notin W_\lambda \mid X_0 = w_i)\\
&= \sum_{u_r \in W_\lambda} \mathbb{P}(X_\gamma = u_r, X_{\gamma-1} \notin W_\lambda, \ldots, X_1 \notin W_\lambda \mid X_0 = w_i)\\
&= \sum_{u_r \in W_\lambda} \sum_{u_s,u_n \notin W_\lambda,} \mathbb{P}(X_\gamma = u_r, X_{\gamma-1} = u_s, \ldots, X_1 = u_n \mid X_0 = w_i)\\
&= \ \ldots \ = \sum_{u_r \in W_\lambda} \sum_{u_s \notin W_\lambda} \sum_{u_n \notin W_\lambda} \mathbb{P}(X_\gamma = u_r \mid X_{\gamma-1} = u_s)\\
&\quad \mathbb{P}(X_{\gamma-1} = u_s, X_{\gamma-2} \notin W_\lambda, \ldots, X_2 \notin W_\lambda \mid X_1 = u_n)\\
&\quad \mathbb{P}(X_1 = u_n \mid X_0 = w_i)\\
&= e_{\ell^k;r_i}\ \tilde{P}_{\Omega^k W_\lambda^c}\ [\tilde{P}_{W_\lambda^c W_\lambda^c}]^{\gamma-2}\tilde{P}_{W_\lambda^c W_\lambda}\ \mathbf{1}'_\lambda.
\end{aligned}
$$

References

[Barbu *et al.*, 2004]Barbu, V., Boussemart, M., Limnios, N.,(2004), Discrete time semi-Markov processes for reliability and survival analysis, *Communications in Statistics - Theory and Methods*, 33(11).

[Biggins and Cannings, 1987]Biggins J D. , Cannings C.,(1987),Markov renewal processes, counters and repeated sequences in markov chains. *Advanced Applied probability trust 19*, 521-545.

[Blom and Thorburn, 1982]Blom and Thorburn,(1982) How many random digits are required until given sequences are obtained, *Applied probability trust*.

[Chryssaphinou and Papastavridis, 1990]Chryssaphinou. O., Papastavridis S., (1990), The occurrence of sequence patterns in repeated dependent experiments,*Th. Probab. Appl. 35* , 145-152.

[Feller, 1968]Feller W.,(1968) An introduction to Probability Theory and its Applications, Wiley, New York.

[Guibas and Odlyzko, 1981]Guibas and Odlyzko ,(1981), String Overlaps, pattern matching and nontransitive games, *Journal of combinatorial Theory*.

[Pittenger, 1987]Pittenger, (1987) Hitting times of sequences , *Stochastic processes and their applications* 24, 225-240 North Holland.

[Robin and Daudin, 1999]Robin S., Daudin J.J., (1999), Exact distribution of word occurrences in a random sequence of letters, *Journal of applied probability* , Vol 36, 179-193.

[Robin and Daudin, 2001]Robin S., Daudin J.J., (2001) Exact distributions of the distances between any occurrence of a set of words, *Annals of the Institute of Statistical Mathematics* , Vol 36 (4) 895-905.