OTHER RESULTS ON THE MARKOVIAN INEQUALITY $P(X_s \in A_s, X_{s+1} \in A_{s+1}, ..., X_t \in A_t) \leq \bar{\alpha}(Q_{s,t})$

UDREA PĂUN

Communicated by Marius Iosifescu

Let $(X_n)_{n\geq 0}$ be a finite Markov chain with state space S and transition matrices $(P_n)_{n\geq 1}$. (The case when S is countable can be considered similarly.) Let $0 \leq s < t$ $(s, t \in \mathbf{N})$. Let $A_s, A_{s+1}, ..., A_t \subseteq S$, $A_s, A_{s+1}, ..., A_t \neq \emptyset, S$. In [U. Păun, $P(X_s \in A_s, X_{s+1} \in A_{s+1}, ..., X_t \in A_t)$ *in the Markov chain case*: *from an upper bound to a method*, Rev. Roumaine Math. Pures Appl. **57** (2012), 145–158] was shown that

$$P(X_s \in A_s, X_{s+1} \in A_{s+1}, ..., X_t \in A_t) \leq \bar{\alpha}(Q_{s,t}),$$

where

$$\bar{\alpha}(P) := \frac{1}{2} \max_{1\leq i,j\leq m} \sum_{k=1}^{n} |P_{ik} - P_{jk}|$$

for any stochastic $m \times n$ matrix P ($\bar{\alpha} = 1 - \alpha$, α is the Dobrushin ergodicity coefficient) and $Q_{s,t} := Q_{s+1}Q_{s+2}...Q_t$, $Q_{s+1}, Q_{s+2}, ..., Q_t$ are matrices which depend on (A_s, A_{s+1}) and $P_{s+1}, (A_{s+1}, A_{s+2})$ and $P_{s+2}, ..., (A_{t-1}, A_t)$ and P_t, respectively. In this article, we investigate some special cases for which the above inequality is even an equation – among other things, we use some new results on the monotone Markov chains. This investigation leads, in particular, to equations for the reliability of certain systems (more generally, for the cumulative distribution function of certain random variables) and for certain probabilities on certain random variables from waiting time random variable theory. Based on these equations, we give several results, such as submultiplicative properties, a more general property than the lack-of-memory one of random variables with geometric distribution, and equations for expectation. So, we have a new and fruitful method to obtain basic results for runs, patterns, etc. This method, based on ergodicity coefficients, etc., can work, on certain matters, in many cases where the classical methods from probability theory hardly work or do not work.

AMS 2010 Subject Classification: 60J10, 60J20, 60E15, 60E99.

Key words: Markov chain, ergodicity coefficient, monotone Markov chain, Markov chain method, reliability, waiting time random variable, equation, inequality, expectation, upper bound having the lack-of-memory property.

1. SOME BASIC RESULTS

In this section, we give some results from finite Markov chain theory. Other results – on the Markov chains as well – are given in Section 2.

In this article, a vector is a row vector and a stochastic matrix is a row stochastic matrix.

Set

$$\langle m\rangle = \{1,2,...,m\} \ \ (m \geq 1),$$

$$S_{m,n} = \{P \mid P \text{ is a stochastic } m\times n \text{ matrix}\},$$

$$N_{m,n} = \{P \mid P \text{ is a nonnegative } m\times n \text{ matrix}\},$$

$$S_m = S_{m,m},$$

$$N_m = N_{m,m}.$$

The entry (i,j) of a matrix Z will be denoted Z_{ij} or, if confusion can arise, $Z_{i\to j}$.

Consider a finite Markov chain $(X_n)_{n\geq 0}$ with state space $S=\langle r\rangle$ and transition matrices $(P_n)_{n\geq 1}$. (We use $S=\langle r\rangle$ for simplification; S can be any finite set.) We shall also refer to it as the (finite) Markov chain $(P_n)_{n\geq 1}$ (with state space $S=\langle r\rangle$). Set

$$P_{m,n} = P_{m+1}P_{m+2}...P_n = \left((P_{m,n})_{ij}\right)_{i,j\in S}, \forall m,n, 0\leq m<n.$$

Let $P=(P_{ij})\in N_{m,n}$. Let $\emptyset \neq U \subseteq \langle m\rangle$ and $\emptyset \neq V \subseteq \langle n\rangle$. Set

$$P_U = (P_{ij})_{i\in U, j\in\langle n\rangle},\ P^V = (P_{ij})_{i\in\langle m\rangle, j\in V},\ P_U^V = (P_{ij})_{i\in U, j\in V}$$

(P_U, P^V, and P_U^V are matrices; *e.g.*, if

$$P = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix},$$

then, *e.g.*,

$$P_{\{2\}} = \begin{pmatrix} 4 & 5 & 6 \end{pmatrix}, P^{\{3\}} = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}, \text{ and } P_{\{1,2\}}^{\{2\}} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}).$$

Set

$$\alpha(P) = \min_{1\leq i,j\leq m} \sum_{k=1}^{n} \min(P_{ik}, P_{jk}),$$

$$\bar{\alpha}(P) = \frac{1}{2} \max_{1\leq i,j\leq m} \sum_{k=1}^{n} |P_{ik} - P_{jk}|.$$

If $P \in S_{m,n}$, then $\alpha(P)$ is called the *Dobrushin ergodicity coefficient of* P (see, *e.g.*, [3] or [8, p. 56]). Set

$$\mu(P) = \max_{j \in \langle n \rangle} \min_{i \in \langle m \rangle} P_{ij},$$

$$\bar{\mu}(P) = 1 - \mu(P).$$

If $P \in S_{m,n}$, then $\mu(P)$ is called the *Markov ergodicity coefficient of* P (see, *e.g.*, [8, p. 56]).

THEOREM 1.1. (i) $\bar{\alpha}(P) = 1 - \alpha(P), \forall P \in S_{m,n}$.

(ii) $\|\nu P - \xi P\|_1 \leq \|\nu - \xi\|_1 \bar{\alpha}(P), \forall \nu, \xi,$ *ν and ξ are probability distributions on* $\langle m \rangle, \forall P \in S_{m,n}$.

(iii) $\bar{\alpha}(PQ) \leq \bar{\alpha}(P)\bar{\alpha}(Q), \forall P \in S_{m,n}, \forall Q \in S_{n,p}$.

Proof. (i) See, *e.g.*, [8, p. 57] or [9, p. 144].

(ii) See, *e.g.*, [3] or [9, p. 147].

(iii) See, *e.g.*, [3], or [8, pp. 58–59], or [9, p. 145]. □

Let $(X_n)_{n \geq 0}$ be a Markov chain with state space $S = \langle r \rangle$. (The case when S is countable can be considered similarly.) Let $0 \leq s < t$ $(s, t \in \mathbf{N})$. Let $A_s, A_{s+1}, ..., A_t \subseteq S$, $A_s, A_{s+1}, ..., A_t \neq \emptyset, S$. Consider the fictive states $\overline{s}, \overline{s+1}, ..., \overline{t}$ ($\overline{s}, \overline{s+1}, ..., \overline{t} \notin S$). Set

$$Q_u = \left((Q_u)_{ij}\right)_{i \in A_{u-1} \cup \left\{\overline{u-1}\right\}, j \in A_u \cup \{\overline{u}\}},$$

$$(Q_u)_{ij} = \begin{cases} (P_u)_{ij} & \text{if } i \in A_{u-1}, j \in A_u, \\ 1 - \sum\limits_{k \in A_u} (P_u)_{ik} & \text{if } i \in A_{u-1}, j = \overline{u}, \\ 0 & \text{if } i = \overline{u-1}, j \in A_u, \\ 1 & \text{if } i = \overline{u-1}, j = \overline{u}, \end{cases}$$

$\forall u \in \{s+1, s+2, ..., t\}, \forall i \in A_{u-1} \cup \left\{\overline{u-1}\right\}, \forall j \in A_u \cup \{\overline{u}\}$; we consider that $\left(\overline{s}, \overline{s+1}\right), \left(\overline{s+1}, \overline{s+2}\right), ..., \left(\overline{t-1}, \overline{t}\right)$ are the last entries of $Q_{s+1}, Q_{s+2}, \ldots, Q_t$, respectively.

When $|S - A_u| = 1$, where $u \in \{s, s+1, ..., t\}$, supposing that $S - A_u = \{i_u\}$, we can work, if we want, with the state i_u of the chain instead of the fictive state $\overline{u}$, see, *e.g.*, Theorems 1.4 and 1.5 and their proofs, see, *e.g.*, also the proofs of Theorems 4.10 and 4.15.

THEOREM 1.2 ([16]). *Under the above conditions we have*

$$P(X_s \in A_s, X_{s+1} \in A_{s+1}, ..., X_t \in A_t) \leq \bar{\alpha}(Q_{s,t})$$

$(Q_{s,t} := Q_{s+1} Q_{s+2} ... Q_t)$.

Proof. See [16]. (The proof is based on Theorem 1.1(ii) and other results; obviously, this proof depends on the fact that – our choice – $\left(\overline{s}, \overline{s+1}\right)$, $\left(\overline{s+1}, \overline{s+2}\right)$, ..., $\left(\overline{t-1}, \overline{t}\right)$ are the last entries of $Q_{s+1}, Q_{s+2}, ..., Q_t$, respectively.) □

Definition 1.3. Let $(P_n)_{n\geq 1}$ be a Markov chain with state space $S = \langle r \rangle$. A state $i \in S$ is called *absorbing* if $(P_n)_{ii} = 1, \forall n \geq 1$.

Below we give an important special case – Theorem 1.4 when $r \geq 2$ – of Theorem 1.2.

THEOREM 1.4 ([16]). *Let $(X_n)_{n\geq 0}$be a Markov chain with state space $S = \langle r \rangle$ and transition matrices $(P_n)_{n\geq 1}$. Suppose that r is an absorbing state. Then*

$$P\left(X_n < r\right) \leq \bar{\alpha}\left(P_{0,n}\right), \forall n \geq 1$$

(this inequality also holds for $n = 0$ if we set $P_{0,0} = I_r$).

Proof. See [16]. □

For applications of Theorem 1.4 to reliability theory, see [16] and, here, Section 3. Section 4 contains applications of this result to waiting time random variable theory.

Another important special case of Theorem 1.2 is the next result when $r \geq 2$.

THEOREM 1.5 ([17]). *Let $(X_n)_{n\geq 0}$be a Markov chain with state space $S = \langle r \rangle$ and transition matrices $(P_n)_{n\geq 1}$. Suppose that r is an absorbing state. Then*

$$P\left(X_0 < r, X_1 < r, ..., X_{n-1} < r, X_n = r\right) \leq \bar{\alpha}\left(P_{0,n-1}Q_n\right), \forall n \geq 1$$

($P_{0,0} := I_r$; $Q_n := I_1$ if $r = 1, \forall n \geq 1$).

Proof. See [17] (we can work with the state r instead of the fictive states $\overline{0}, \overline{1}, ..., \overline{n-1}$ – we only need the fictive state $\overline{n}$ (for $r = 2$, we can work, if we want, with the state 1 of the chain instead of the fictive state $\overline{n}$); in this case, (*i.e.*, when we work with the state r instead of the fictive states $\overline{0}, \overline{1}, ..., \overline{n-1}$), $Q_l = P_l$, $\forall l \in \langle n-1 \rangle$). □

For applications of Theorem 1.5 to waiting time random variable theory, see [17] and, here, Section 4.

Set

$$\mathrm{Par}\left(E\right) = \left\{\Delta \mid \Delta \text{ is a partition of } E\right\},$$

where E is a nonempty set. We shall agree that the partitions do not contain the empty set.

Set

$$\left(\{i\}\right)_{i\in\{s_1,s_2,...,s_t\}} = \left(\{s_1\}, \{s_2\}, ..., \{s_t\}\right);$$

$$(\{i\})_{i\in\{s_1,s_2,...,s_t\}} \in \operatorname{Par}(\{s_1, s_2, ..., s_t\}).$$

E.g., $(\{i\})_{i\in\langle 10\rangle} = (\{1\}, \{2\}, ..., \{10\}) \in \operatorname{Par}(\langle 10\rangle)$.

Below we give part of Theorem 1.8 from [15] and this in the stochastic case only, see also [16, Theorem 2.1] or [17, Theorem 1.6].

THEOREM 1.6. *Let* $P_1 \in S_{m_1,m_2}$, $P_2 \in S_{m_2,m_3}, ..., P_n \in S_{m_n,m_{n+1}}$. *Let* $\Delta_1 = (\langle m_1\rangle)$, $\Delta_2 \in Par(\langle m_2\rangle), ..., \Delta_n \in Par(\langle m_n\rangle)$, $\Delta_{n+1} = (\{i\})_{i\in\langle m_{n+1}\rangle}$. *Consider the matrices* $L_l = ((L_l)_{VW})_{V\in\Delta_l, W\in\Delta_{l+1}}$, $l \in \langle n\rangle$ ($(L_l)_{VW}$ *is the entry* (V, W) *of matrix* L_l), *where*

$$(L_l)_{VW} := \min_{i\in V} \sum_{j\in W} (P_l)_{ij}, \forall l \in \langle n\rangle, \forall V \in \Delta_l, \forall W \in \Delta_{l+1}.$$

Then

$$\alpha(P_1P_2...P_n) \geq \sum_{K\in\Delta_{n+1}} (L_1L_2...L_n)_{\langle m_1\rangle K}.$$

(*Since* $L_1L_2...L_n$ *is an* $1 \times |\langle m_{n+1}\rangle|$ *matrix, it can be thought of as a row vector, but above we used and below we shall use, if necessary, the matrix notation for its entries instead of the vector one. Above the matrix notation* $(L_1L_2...L_n)_{\langle m_1\rangle K}$ *was used instead of the vector one* $(L_1L_2...L_n)_K$ *because, in this article, the notation* A_U, *where* $A \in N_{p,q}$ *and* $\emptyset \neq U \subseteq \langle p\rangle$, *means something different.*)

Proof. See [15]. □

Set

$$\mathcal{R}^{ij}_{m,n} = \{P \mid P \in S_{m,n} \text{ and } P_{ij} = 1\},$$

$$\mathcal{R}^{ij}_{m} = \mathcal{R}^{ij}_{m,m},$$

$$\mathcal{R}^{i}_{m} = \mathcal{R}^{ii}_{m}$$

(see [14] for $\mathcal{R}^{ij}_{m,n}$ and $\mathcal{R}^{i}_{m}$).

THEOREM 1.7 ([14]). *Let* $P \in \mathcal{R}^{ij}_{m,n}$. *Then*

$$\alpha(P) = \mu(P) = \mu\left(P^{\{j\}}\right)$$

($\mu\left(P^{\{j\}}\right) = \min_{k\in\langle m\rangle} P_{kj}$).

Proof. See [14]. □

Theorem 1.7 is important because it gives an easier way to compute $\alpha(P)$ and (see Theorem 1.1(i)) $\bar{\alpha}(P)$ when $P \in \mathcal{R}^{ij}_{m,n}$. On the other hand, this result can be used to obtain properties of random variables, see the next sections.

2. MONOTONE MARKOV CHAINS

In this section, we give some results on the monotone Markov chains (see, *e.g.*, also [2] and [10–11]).

Definition 2.1 (see, *e.g.*, [10, p. 164]). Let $x, y \in \mathbf{R}^n$ be two (row) stochastic vectors. We say that y is *larger stochastically than* x or that y *dominates* x if

$$\sum_{k=l}^{n} x_k \leq \sum_{k=l}^{n} y_k, \ \forall l \in \langle n \rangle .$$

Set $x \underline{<<} y$ when y is larger stochastically than x.

Definition 2.2 (see, *e.g.*, [10, p. 165] for square matrices). Let $P \in S_{m,n}$. We say that P is a (*stochastically*) *monotone matrix* if

$$P_{\{1\}} \underline{<<} P_{\{2\}} \underline{<<} \ldots \underline{<<} P_{\{m\}}$$

($P_{\{1\}}, P_{\{2\}}, \ldots, P_{\{m\}}$ are the first, the second, ..., the mth row of P, respectively).

Remark 2.3. By permutation of rows and columns, some nonmonotone matrices can be transformed in monotone ones. *E.g.*,

$$P = \begin{pmatrix} \frac{9}{10} & 0 & \frac{1}{10} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

is not a monotone matrix. Setting

$$U = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

(U is a permutation matrix), the matrix

$$Q := UPU' = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{10} & \frac{9}{10} \end{pmatrix}$$

is a monotone matrix, where U' is the transpose of U. (See also [2, the example from p. 309].)

Set

$$\mathcal{M}_{m,n} = \{P \mid P \text{ is a monotone } m \times n \text{ matrix}\}$$

and

$$\mathcal{M}_m = \mathcal{M}_{m,m}$$

(see, *e.g.*, [10, p. 165] for $\mathcal{M}_m$).

THEOREM 2.4. *Let* $P \in \mathcal{M}_{m,n}$ *and* $Q \in \mathcal{M}_{n,p}$. *Then* $PQ \in \mathcal{M}_{m,p}$.

Proof. For the special case $m = n = p$, see, *e.g.*, [10, p. 166] or [11]. Now, we consider the general case. We construct the stochastic matrices

$$A = \begin{pmatrix} P & 0_{m\times(m+p-1)} & 0_{m\times 1} \\ 0_{(n+p)\times n} & 0_{(n+p)\times(m+p-1)} & e'(n+p) \end{pmatrix} \in S_{m+n+p}$$

and

$$B = \begin{pmatrix} Q & 0_{n\times(m+n-1)} & 0_{n\times 1} \\ 0_{(m+p)\times p} & 0_{(m+p)\times(m+n-1)} & e'(m+p) \end{pmatrix} \in S_{m+n+p},$$

where $0_{s\times t}$ is the zero $s \times t$ matrix, $\forall s, t \geq 1$, and $e'(t)$ is the transpose of $e(t) := (1, 1, ..., 1) \in \mathbf{R}^t, \forall t \geq 1$. Obviously, A and B are monotone matrices. It follows from the above special case that AB is monotone. Since

$$AB = \begin{pmatrix} PQ & 0_{m\times(m+n-1)} & 0_{m\times 1} \\ 0_{(n+p)\times p} & 0_{(n+p)\times(m+n-1)} & e'(n+p) \end{pmatrix}$$

($AB \in S_{m+n+p}$) and

$$(AB)_{\{1\}} \underline{<<} (AB)_{\{2\}} \underline{<<} ... \underline{<<} (AB)_{\{m\}},$$

we have

$$(PQ)_{\{1\}} \underline{<<} (PQ)_{\{2\}} \underline{<<} ... \underline{<<} (PQ)_{\{m\}}.$$

Therefore, PQ is monotone. □

THEOREM 2.5. *Let* $P \in \mathcal{R}^{mn}_{m,n} \bigcap \mathcal{M}_{m,n}$. *Then*

$$\alpha(P) = \mu(P) = \mu\left(P^{\{n\}}\right) = P_{1n}.$$

Proof. By Theorem 1.7,

$$\alpha(P) = \mu(P) = \mu\left(P^{\{n\}}\right).$$

We prove that

$$\mu\left(P^{\{n\}}\right) = P_{1n}.$$

We have

$$\mu\left(P^{\{n\}}\right) = \min_{1\leq k\leq m} P_{kn}.$$

On the other hand, since $P \in \mathcal{M}_{m,n}$, we have

$$P_{1n} \leq P_{2n} \leq P_{3n} \leq ... \leq P_{mn} = 1.$$

Consequently,

$$\mu\left(P^{\{n\}}\right) = P_{1n}. \quad \square$$

Definition 2.6 (see, *e.g.*, [10, p. 168] for the homogeneous case and constant state space). Let $(P_n)_{n\geq 1}$ be a (finite) Markov chain with state space $S = \langle r \rangle$ or, more generally, with time varying (finite) state space, *i.e.*, instead of S we have a sequence of state spaces $S_0 = \langle r_0 \rangle, S_1 = \langle r_1 \rangle, \ldots$ $(r_0, r_1, \ldots \geq 1)$, see, *e.g.*, [8, p. 215]. We say that the chain is *monotone* if P_n is a monotone matrix, $\forall n \geq 1$.

Remark 2.7 (See Remark 2.3 again). When we construct a Markov chain, an auspicious labeling of states could lead to a monotone one.

THEOREM 2.8. *Let $(X_n)_{n\geq 0}$ be a monotone Markov chain with state space $S = \langle r \rangle$ or, more generally, with state spaces $S_0 = \langle r_0 \rangle, S_1 = \langle r_1 \rangle, \ldots$ and transition matrices $(P_n)_{n\geq 1}$. Suppose that the initial (probability) distribution is $p_0 = (1, 0, \ldots., 0)$ and r is an absorbing state or, more generally, $(P_1)_{r_0 r_1} = 1$, $(P_2)_{r_1 r_2} = 1, \ldots$ (in this case, we call $(r_n)_{n\geq 0}$ the absorbing sequence). Then*

$$P(X_n = r) = \alpha(P_{0,n}) = \mu(P_{0,n}) = \mu\left((P_{0,n})^{\{r\}}\right) = (P_{0,n})_{1r}, \forall n \geq 1,$$

or, more generally,

$$P(X_n = r_n) = \alpha(P_{0,n}) = \mu(P_{0,n}) = \mu\left((P_{0,n})^{\{r_n\}}\right) = (P_{0,n})_{1r_n}, \forall n \geq 1,$$

and, therefore,

$$P(X_n < r) = \bar{\alpha}(P_{0,n}) = \bar{\mu}(P_{0,n}) = \bar{\mu}\left((P_{0,n})^{\{r\}}\right) = 1 - (P_{0,n})_{1r}, \forall n \geq 1,$$

or, more generally,

$$P(X_n < r_n) = \bar{\alpha}(P_{0,n}) = \bar{\mu}(P_{0,n}) = \bar{\mu}\left((P_{0,n})^{\{r_n\}}\right) = 1 - (P_{0,n})_{1r_n}, \forall n \geq 1.$$

Proof. It follows from Theorems 1.1(i), 2.4, and 2.5 and the fact that

$$P(X_n = r) = (p_0 P_{0,n})_r = (P_{0,n})_{1r}, \forall n \geq 1,$$

or, more generally,

$$P(X_n = r_n) = (p_0 P_{0,n})_{r_n} = (P_{0,n})_{1r_n}, \forall n \geq 1. \qquad \square$$

Theorem 2.8 gives one way in which the inequality from Theorem 1.4 becomes an equation. Another way will be given in Section 4.

3. APPLICATIONS TO RELIABILITY THEORY

In this section, we give equations and a submultiplicative property for the reliability of certain systems with independent components. More generally, these results refer to the cumulative distribution function of certain random variables – it is easy to see which are those random variables.

Set

$$\langle\langle m \rangle\rangle = \{0, 1, ..., m\} \ \ (m \geq 0)$$

and

$$\text{supp } \nu = \{i \mid i \in W \text{ and } \nu_i > 0\}$$

(the support of ν), where W is a finite nonempty set and $\nu = (\nu_i)_{i \in W}$ is a probability distribution on W.

Recall that the entry (i, j) of a matrix Z is denoted Z_{ij} or, if confusion can arise, $Z_{i \to j}$.

Definition 3.1 ([16]). A Markov chain with state space $S = \langle\langle k \rangle\rangle$, where $k \geq 1$ $(k \in \mathbf{N})$, initial (probability) distribution ψ_0 with supp $\psi_0 \subseteq \langle\langle k-1 \rangle\rangle$, and transition matrices

$$P_n = \begin{array}{c} 0 \\ 1 \\ \vdots \\ \\ k - w_n \\ \vdots \\ k-1 \\ k \end{array} \left(\begin{array}{ccccccc} p_n & \cdots & & q_n & & & \\ & p_n & \cdots & & q_n & & \\ & & \ddots & & & \ddots & \\ & & & & & & \\ & & & & p_n & \cdots & q_n \\ & & & & & \ddots & \vdots \\ & & & & & p_n & q_n \\ & & & & & & 1 \end{array} \right), n \geq 1,$$

(the columns are labeled similarly, *i.e.*, $0, 1, ..., k$ from left to right) with $(P_n)_{i \to i+w_n} = q_n, \forall n \geq 1, \forall i \in S, i + w_n \leq k$, and $(P_n)_{ik} = q_n, \forall n \geq 1, \forall i \in S - \{k\}, i + w_n > k$, where w_n is a natural number, $1 \leq w_n \leq k, \forall n \geq 1$ (obviously, $(P_n)_{ii} = p_n, \forall n \geq 1, \forall i \in \langle\langle k-1 \rangle\rangle$, and $(P_n)_{kk} = 1, \forall n \geq 1$ (*i.e.*, k is an absorbing state)), is called *weighted k-out-of-∞: F*. A weighted k-out-of-∞: F Markov chain with $w_n = 1, \forall n \geq 1$, is called *$k$-out-of-$\infty$: F*. A 1-out-of-$\infty$: F Markov chain $(k = 1)$ is called *series*. We call w_n the *weight of* $P_n, \forall n \geq 1$.

Consider a weighted v-component system $(v \geq 1)$, *i.e.*, a system with v components, the component n having a weight, say, $w_n, \forall n \in \langle v \rangle$. We only work with independent components. Suppose that $w_n \geq 1$ and $w_n \in \mathbf{N}$. The component n fails with probability, say, $q_n, \forall n \in \langle v \rangle$. A weighted k-out-of-v: F system is a weighted v-component system which fails if and only if the total weight of failed components is at least k (see, *e.g.*, [12] and [13, p. 279]). Following the Markov chain method (see [4]; see, *e.g.*, also [1, pp. 13–14], [6–7], [12–13], and [16–17]), this system determines v stochastic matrices, say,

$P_1, P_2, ..., P_v$ (we associate a stochastic matrix with each component of the system), where P_n is identical with P_n from Definition 3.1, $\forall n \in \langle v \rangle$ (since P_n, $n \in \langle v \rangle$, from Definition 3.1 are stochastic matrices, we have, there and here, $p_n = 1 - q_n, \forall n \in \langle v \rangle$). To work with Markov chains, since the matrices $P_1, P_2, ..., P_v$ do not determine a Markov chain (not even when an initial distribution is given), we consider a weighted k-out-of-∞: F Markov chain, $(X_n)_{n \geq 0}$, having the first v matrices even these ones ("∞" from weighted k-out-of-∞: F was suggested by the fact that any chain has an infinite number of transition matrices (obviously, it is possible as some of them or even all be identical)). Further, using this weighted k-out-of-∞: F Markov chain framework and the fact that the reliability of a v-component system, R_v, is the probability that this works, it follows that the reliability of above weighted k-out-of-v: F system, $R_v = R_v (k, q_1, q_2, ..., q_v, w_1, w_2, ..., w_v)$, is equal to $P(X_v < k)$, *i.e.*,

$$R_v = P(X_v < k).$$

Consequently, to give inequalities and equations for R_v, we can work in the weighted k-out-of-∞: F Markov chain framework, see also [16].

Consider a weighted k-out-of-v: F system (recall that its components are independent). We associate this system with a weighted k-out-of-∞: F Markov chain, $(X_n)_{n \geq 0}$, as above. (Obviously, this association is not unique, but this fact does not count.) If $\psi_0 = (1, 0, ..., 0)$ is the initial distribution of chain (this is the usual case), then, by Theorem 2.8, since $(X_n)_{n \geq 0}$ is monotone, we have other equations for the reliability R_v of this system in the next result.

THEOREM 3.2. *Under the above conditions we have*

$$R_v = \bar{\alpha}(P_{0,v}) = \bar{\mu}(P_{0,v}) = \bar{\mu}\left((P_{0,v})^{\{k\}}\right) = 1 - (P_{0,v})_{0k}.$$

Proof. See above. □

Remark 3.3. The equation $R_v = \bar{\alpha}(P_{0,v})$ from Theorem 3.2 is important for two reasons. First, using Theorems 1.1, 1.4, and 1.6 only, we obtain (see also [16])

$$R_v \leq \bar{\alpha}(P_{0,v}) \leq C_v$$

for any upper bound C_v of $\bar{\alpha}(P_{0,v})$ we find by Theorems 1.1 and 1.6 – consequently, we have

$$R_v < \bar{\alpha}(P_{0,v}) \leq C_v$$

(this is the worst possible case because if it holds, then the best upper bound C_v of R_v we could obtain is equal to $\bar{\alpha}(P_{0,v})$) or

$$R_v = \bar{\alpha}(P_{0,v}) \leq C_v$$

(this is the best possible case because if it holds, then the best upper bound C_v of R_v we could obtain is even equal to R_v). By Theorem 3.2 the latter case holds here. Second, this equation can be used to obtain other properties of the reliability for weighted k-out-of-v: F systems, $k, v \geq 1$. *E.g.*, if we consider $v = v_1 + v_2$, $v_1, v_2 \geq 1$, and the subsystems weighted k-out-of-v_1: F and weighted k-out-of-v_2: F of weighted k-out-of-v: F system with the reliabilities R_{v_1} and R_{v_2}, respectively, and, moreover, suppose that the components of the first subsystem are the first v_1 components of weighted k-out-of-v: F system and the associate chains of subsystems have both the initial distribution $(1, 0, ..., 0)$, then

$$R_v \leq R_{v_1} R_{v_2},$$

i.e., the reliability for weighted k-out-of-v: F systems, k is fixed, $v \geq 1$ – here, thought of as a function $v \longmapsto R_v$ $(v \geq 1)$ – is submultiplicative. (This property is important because, *e.g.*, if we know R_{v_1} and R_{v_2} or upper bounds of theirs, then we can obtain upper bounds for R_v.) Indeed, by Theorem 1.1(iii), we have

$$R_v = \bar{\alpha}\left(P_{0,v}\right) = \bar{\alpha}\left(P_{0,v_1+v_2}\right) = \bar{\alpha}\left(P_{0,v_1} P_{v_1,v_1+v_2}\right) \leq$$

$$\leq \bar{\alpha}\left(P_{0,v_1}\right) \bar{\alpha}\left(P_{v_1,v_1+v_2}\right) = R_{v_1} R_{v_2}.$$

(Note that we are in a happy situation here because we can give another proof of the inequality $R_v \leq R_{v_1} R_{v_2}$ as follows. We associate component n of the weighted k-out-of-v: F system with the random variable Y_n with

$$Y_n = \begin{cases} 0 & \text{if the component } n \text{ works,} \\ 1 & \text{if the component } n \text{ does not work,} \end{cases}$$

$\forall n \in \langle v \rangle$. Consequently, $P\left(Y_n = 0\right) = p_n$ and $P\left(Y_n = 1\right) = q_n, \forall n \in \langle v \rangle$, and $Y_1, Y_2, ..., Y_v$ are independent because the components of system are independent. Set

$$X = w_1 Y_1 + w_2 Y_2 + ... + w_v Y_v,$$

$$Y = w_1 Y_1 + w_2 Y_2 + ... + w_{v_1} Y_{v_1},$$

$$Z = w_{v_1+1} Y_{v_1+1} + w_{v_1+2} Y_{v_1+2} + ... + w_v Y_v.$$

By the way, X, Y, Z are random variables with binomial distribution when $w_1 = w_2 = ... = w_v = 1$. Since $Y_1, Y_2, ..., Y_v$ are independent, it follows that Y and Z are independent. Finally,

$$R_v = P\left(X < k\right) = P\left(Y + Z < k\right) \leq$$

$$\leq P\left(\{Y < k\} \cap \{Z < k\}\right) = P\left(Y < k\right) P\left(Z < k\right) = R_{v_1} R_{v_2}.\ )$$

Other examples of monotone Markov chains from reliability theory are the weighted consecutive-k-out-of-∞: F Markov chains (these contain the consecutive-k-out-of-∞: F Markov chains) and, more generally, the weighted m-consecutive-k-out-of-∞: F Markov chains (these contain the m-consecutive-k-out-of-∞: F Markov chains), see [14] for the definitions of these chains, see also [16]. The weighted consecutive-k-out-of-v: F systems (see, *e.g.*, [12] for their definition) are associated with the weighted consecutive-k-out-of-∞: F Markov chains while the weighted m-consecutive-k-out-of-v: F systems (see, *e.g.*, [12] for their definition) are associated with the weighted m-consecutive-k-out-of-∞: F Markov chains. These associations are similar to that one between the weighted k-out-of-v: F systems and weighted k-out-of-∞: F Markov chains (see Definition 3.1 and after it again). For each v-component system, which is weighted consecutive-k-out-of-v: F or, more generally, weighted m-consecutive-k-out-of-v: F, the above equations and inequalities on its reliability (see Theorem 3.2 and Remark 3.3 again) hold as well – do these hold for any v-component system with independent components? or for a large class of v-component systems with independent components? (see also Section 4, Results based on induction)

4. APPLICATIONS TO WAITING TIME RANDOM VARIABLE THEORY

In this section, we give equations for $P(X = n)$ and $P(X > n)$ of certain waiting time random variables. Using these equations, we obtain, in particular, properties of the waiting time random variables, such as a more general property than the lack-of-memory one of random variables with geometric distribution. In this section, the trials are only independent – they are or not identically distributed.

A. Results for $P(X > n)$.

A.1. Results based on monotone chains. We give two examples. For these examples, A.1.1 and A.1.2, we consider that the possible outcomes for each trial are s ("success") and f ("failure"). Set

$$p_n = \text{the probability that } s \text{ occur in the } n\text{th trial}$$

and

$$q_n = \text{the probability that } f \text{ occur in the } n\text{th trial}, \forall n \geq 1$$

$(p_n, q_n \geq 0,\ p_n + q_n = 1, \forall n \geq 1)$.

A.1.1. X = the waiting time of kth occurrence of s ($k \geq 1$; see, *e.g.*, [1, p. 103]; see also [5, pp. 164–167]; X is a random variable with negative binomial distribution of order k when trials – we work with independent trials only – are identically distributed).

Following the Markov chain method (recall [4]; recall, *e.g.*, also [1, pp. 13–14], [6–7], [12–13], and [16–17]) we associate a stochastic matrix with each trial; the Markov chain corresponding to X above is, say, $(X_n)_{n\geq 0}$ with the state space $S = \langle\langle k\rangle\rangle$ $(k \geq 1)$, initial (probability) distribution $\psi_0 = (1, 0, ..., 0)$, where $P(X_0 = 0) = 1$ (we can take, in particular, $X_0 = 0$), and transition matrices

$$P_n = \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ k-1 \\ k \end{array} \begin{pmatrix} q_n & p_n & & & & \\ & q_n & p_n & & & \\ & & q_n & p_n & & \\ & & & \ddots & \ddots & \\ & & & & q_n & p_n \\ & & & & & 1 \end{pmatrix}, n \geq 1$$

(the columns are labeled similarly, *i.e.*, $0, 1, 2, ..., k-1, k$ from left to right). It follows that

$$P(X > n) = P(X_n < k), \forall n \geq 0.$$

Below we give other equations for $P(X > n)$.

THEOREM 4.1. *Under the above conditions we have*

$$P(X > n) = \bar{\alpha}(P_{0,n}) = \bar{\mu}(P_{0,n}) = \bar{\mu}\left((P_{0,n})^{\{k\}}\right) = 1 - (P_{0,n})_{0k}, \forall n \geq 0.$$

Proof. It follows from Theorem 2.8 because $(X_n)_{n\geq 0}$ above is a monotone chain. (We can give another proof based on induction instead of monotone chains, see A.2). □

Remark 4.2. We have reasons similar to those from Remark 3.3 on the importance of equation $P(X > n) = \bar{\alpha}(P_{0,n}), \forall n \geq 0$, from Theorem 4.1. (See also the results below.)

THEOREM 4.3 (See also Theorem 2.5(ii) in [17]). *Keeping the conditions from Theorem* 4.1 *we have*

$$E(X) = \sum_{n=0}^{\infty} \bar{\alpha}(P_{0,n})$$

($E(X) =$ *the expectation of* X; $P_{0,0} := I_{k+1}$).

Proof. By a well-known result and Theorem 4.1,

$$E(X) = \sum_{n=1}^{\infty} nP(X = n) = \sum_{n=0}^{\infty} P(X > n) = \sum_{n=0}^{\infty} \bar{\alpha}(P_{0,n}). \quad \square$$

The equation from Theorem 4.3 is a bridge between waiting time random variable theory and Markov chain theory. Both fields can benefit by this connection. *E.g.*, if $E(X) < \infty$, then $\bar{\alpha}(P_{0,n}) \to 0$ as $n \to \infty$, *i.e.*, cf., *e.g.*, [14], the chain $(P_n)_{n\geq 1}$ is strongly ergodic at time 0 (consequently, we need results on the strongly ergodic Markov chains at time 0). Further, the strong ergodicity at time 0 of chain could help to obtain results on the behaviour of X. Suppose now, *e.g.*, that the chain $(P_n)_{n\geq 1}$ is not strongly ergodic at time 0. By the above result, $E(X) = \infty$.

Remark 4.4. The inequality from Theorem 1.1(iii) is even an equation when $m = n = p = 2$ (see, *e.g.*, [8, pp. 58–59] – this follows by direct computation).

Let $m \geq 0$. Let $Y^{(m)}$ be another waiting time random variable defined as follows. The first trial of $Y^{(m)}$ and $(m+1)$th trial of X are identically distributed, the second trial of $Y^{(m)}$ and $(m+2)$th trial of X are also identically distributed, etc. We associate $Y^{(m)}$ with the Markov chain $(Z_n)_{n\geq m}$ with the state space $S = \langle\langle k\rangle\rangle$, initial distribution $\psi_0 = (1, 0, ..., 0)$, where $P(Z_m = 0) = 1$, and transition matrices $V_n = P_n, \forall n > m$, where P_n, $n \geq 1$, are the transition matrices of chain $(X_n)_{n\geq 0}$ associated with X.

THEOREM 4.5. *Under the above conditions we have*

$$P(X > m+n) \leq P(X > m) P\left(Y^{(m)} > n\right), \forall m, n \geq 0,$$

and if, moreover, $k = 1$, we have

$$P(X > m+n) = P(X > m) P\left(Y^{(m)} > n\right), \forall m, n \geq 0.$$

In particular, if trials are identically distributed, we have

$$P(X > m+n) \leq P(X > m) P(X > n), \forall m, n \geq 0,$$

and if, moreover, $k = 1$ (i.e., X is a random variable with geometric distribution), we have

$$P(X > m+n) = P(X > m) P(X > n), \forall m, n \geq 0.$$

Proof. By Theorems 1.1(iii) and 4.1,

$$P(X > m+n) = \bar{\alpha}(P_{0,m+n}) \leq \bar{\alpha}(P_{0,m})\bar{\alpha}(P_{m,m+n}) =$$

$$= P(X > m) P\left(Y^{(m)} > n\right), \forall m, n \geq 0$$

(we set $P_{0,0} = I_{k+1}$). The second inequality follows from the first one because $P\left(Y^{(m)} > n\right) = P(X > n), \forall m, n \geq 0$, when trials are identically distributed. The equations in the special case $k = 1$ follow from Theorem 4.1 and Remark 4.4. $\square$

Set

$$\lfloor x \rfloor = \max \{k \mid k \in \mathbf{Z} \text{ and } k \leq x\},$$

where $x \in \mathbf{R}$.

THEOREM 4.6. *Keeping the conditions before Theorem 4.5, with only the difference that the trials are considered identically distributed, we have*

$$P(X > n_1 + n_2 + ... + n_u) \leq$$
$$\leq P(X > n_1) P(X > n_2) ... P(X > n_u), \forall u \geq 1, \forall n_1, n_2, ..., n_u \geq 0,$$

$$P(X > ht) \leq P(X > t)^h, \forall h, t \geq 0,$$

and

$$P(X > n) \leq \left(1 - p^k\right)^{\lfloor \frac{n}{k} \rfloor}, \forall n \geq 0,$$

where $p := p_1 = p_2 = ...$ *(recall that* $X =$ *the waiting time of* k^{th} *occurrence of* s*).*

Proof. The first inequality follows by Theorem 4.5 and induction. The second one follows by the first. Now, we prove the third inequality.

$$P(X > n) \leq P\left(X > \left\lfloor \frac{n}{k} \right\rfloor k\right) \leq (P(X > k))^{\lfloor \frac{n}{k} \rfloor} =$$
$$= (1 - P(X = k))^{\lfloor \frac{n}{k} \rfloor} = \left(1 - p^k\right)^{\lfloor \frac{n}{k} \rfloor}, \forall n \geq 0. \quad \square$$

The third inequality from Theorem 4.6 is interesting because it gives an upper bound for $P(X > n), \forall n \geq 0$, and information on the speed of convergence of $P(X > n)$ to 0 as $n \to \infty$.

Let $(\Omega, \mathcal{K}, P)$ be a probability space. Let $A, B \in \mathcal{K}$. Set

$$P(A | B) = \begin{cases} \frac{P(A \cap B)}{P(B)} & \text{if } P(B) > 0, \\ 0 & \text{if } P(B) = 0 \end{cases}$$

($P(\cdot | \cdot)$ is the conditional probability).

THEOREM 4.7. *Keeping the conditions before Theorem 4.5 we have*

$$P(X > m + n \mid X > m) \leq P\left(Y^{(m)} > n\right), \forall m, n \geq 0,$$

and if, moreover, $k = 1$, *we have*

$$P(X > m + n \mid X > m) = P\left(Y^{(m)} > n\right), \forall m, n \geq 0$$

(this is a generalization of the lack-of-memory property of random variables with geometric distribution – we call it the lack-of-memory property of X *in the independent case). In particular, if trials are identically distributed, we have*

$$P(X > m + n \mid X > m) \leq P(X > n), \forall m, n \geq 0,$$

and if, moreover, $k=1$ *(i.e., X is a random variable with geometric distribution), we have*

$$P(X>m+n \mid X>m)=P(X>n),\forall m,n\geq 0$$

(this property is known as the lack-of-memory property of X*).*

Proof. By Theorem 4.5,

$$P(X>m+n \mid X>m)=\frac{P(\{X>m+n\}\cap\{X>m\})}{P(X>m)}=$$

$$=\frac{P(X>m+n)}{P(X>m)}\leq\frac{P(X>m)P\left(Y^{(m)}>n\right)}{P(X>m)}=P\left(Y^{(m)}>n\right),$$

$\forall m,n\geq 0$ with $P(X>m)>0$ (the case $P(X>m)=0$ is obvious). The others also follow by Theorem 4.5. □

Definition 4.8. We say that the upper bound $P\left(Y^{(m)}>n\right)$ of conditional probability $P(X>m+n \mid X>m),\forall m,n\geq 0$, from Theorem 4.7, has the *lack-of-memory property in the independent case.* (This definition is based on the fact that $P\left(Y^{(m)}>n\right)$ does not depend on what it happened before the $(m+1)$th trial.) In particular, if trials are identically distributed, then we say that the upper bound $P(X>n)$ of $P(X>m+n \mid X>m),\forall m,n\geq 0$, has the *lack-of-memory property in the independent and identically distributed case.*

A.1.2. $X=$ the waiting time of k consecutive occurrences of s ($k\geq 1$; see, *e.g.*, [6, p. 64]; X is a random variable with geometric distribution of order k when trials are identically distributed).

Following the Markov chain method, the Markov chain corresponding to X above is, say, $(X_n)_{n\geq 0}$ with the state space $S=\langle\langle k\rangle\rangle$ $(k\geq 1)$, initial distribution $\psi_0=(1,0,...,0)$, where $P(X_0=0)=1$ (we can take, in particular, $X_0=0$), and transition matrices

$$P_n=\begin{array}{c} 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ k-1 \\ k \end{array}\left(\begin{array}{ccccccc} q_n & p_n & & & & & \\ q_n & & p_n & & & & \\ \cdot & & & \cdot & & & \\ \cdot & & & & \cdot & & \\ \cdot & & & & & \cdot & \\ q_n & & & & & & p_n \\ & & & & & & 1 \end{array}\right),n\geq 1.$$

It follows that

$$P(X>n)=P(X_n<k),\forall n\geq 0.$$

Since $(X_n)_{n\geq 0}$ is a monotone Markov chain, it follows that all the results from A.1.1 on X from there hold for X above as well.

A.2. Results based on induction. We give two examples.

A.2.1. $X =$ the waiting time of pattern $\Theta = a_{i_1}a_{i_2}...a_{i_k}$ with $a_{i_2}, a_{i_3}, ..., a_{i_k} \neq a_{i_1}$. Suppose that the possible outcomes for each trial are $a_1, a_2, ..., a_m$ ($m \geq 2$) and, obviously, $i_1, i_2, ..., i_k \in \langle m \rangle$ ($k \geq 1$). Consider

$$q_l^{(n)} = \text{the probability that } a_l \text{ occur in the } n\text{th trial}, \forall l \in \langle m \rangle, \forall n \geq 1$$

($q_l^{(n)} \geq 0, \forall l \in \langle m \rangle, \forall n \geq 1$, and $\sum_{l \in \langle m \rangle} q_l^{(n)} = 1, \forall n \geq 1$).

Below we use the empty word from formal language theory. This word, denoted λ, has, as its name suggests, the length (*i.e.*, the number of symbols) equal to 0.

Following the Markov chain method, we associate X with the Markov chain $(X_n)_{n \geq 0}$ with the state space $S = \{\lambda, a_{i_1}, a_{i_1}a_{i_2}, ..., a_{i_1}a_{i_2}...a_{i_{k-1}}, \Theta\}$, initial distribution $\psi_0 = (1, 0, ..., 0)$, where $P(X_0 = \lambda) = 1$ (we can take, in particular, $X_0 = \lambda$), and transition matrices (these are lower Hessenberg matrices, *i.e.*, matrices with zero entries above the first superdiagonal)

$$P_n = \begin{array}{c} \lambda \\ a_{i_1} \\ a_{i_1}a_{i_2} \\ \vdots \\ a_{i_1}a_{i_2}...a_{i_{k-1}} \\ \Theta \end{array} \begin{pmatrix} w_{i_1}^{(n)} & q_{i_1}^{(n)} & & & & \\ w_{i_2}^{(n)} & q_{i_1}^{(n)} & q_{i_2}^{(n)} & & & \\ w_{i_3}^{(n)} & q_{i_1}^{(n)} & & q_{i_3}^{(n)} & & \\ \vdots & \vdots & & & \ddots & \\ w_{i_k}^{(n)} & q_{i_1}^{(n)} & & & & q_{i_k}^{(n)} \\ & & & & & 1 \end{pmatrix}, n \geq 1$$

(recall that the columns are labeled similarly, *i.e.*, λ, a_{i_1}, $a_{i_1}a_{i_2}$, ..., $a_{i_1}a_{i_2}...a_{i_{k-1}}$, Θ from left to right), where

$$w_{i_u}^{(n)} := \begin{cases} 1 - q_{i_1}^{(n)} & \text{if } u = 1, \\ 1 - q_{i_1}^{(n)} - q_{i_u}^{(n)} & \text{if } u \in \{2, 3, ..., k\}, \end{cases}$$

$\forall n \geq 1$. It follows that

$$P(X > n) = P(X_n \neq \Theta), \forall n \geq 0.$$

Below we give other equations for $P(X > n)$ (these are similar to those from Theorem 4.1).

THEOREM 4.9. *Under the above conditions we have*

$$P(X > n) = \bar{\alpha}(P_{0,n}) = \bar{\mu}(P_{0,n}) = \bar{\mu}\left((P_{0,n})^{\{\Theta\}}\right) = 1 - (P_{0,n})_{\lambda\Theta}, \forall n \geq 0$$

$(P_{0,0} := I_{k+1})$.

Proof. By Theorem 1.7,

$$\alpha\left(P_{0,n}\right)=\mu\left(P_{0,n}\right)=\mu\left(\left(P_{0,n}\right)^{\{\Theta\}}\right)=\min_{k\in S}\left(P_{0,n}\right)_{k\Theta},\forall n\geq 0.$$

We prove that

$$\min_{k\in S}\left(P_{0,n}\right)_{k\Theta}=\left(P_{0,n}\right)_{\lambda\Theta},\forall n\geq 0.$$

To see this, first, we prove that

$$\min_{k\in S}\left(P_{s-t,s}\right)_{k\Theta}=\left(P_{s-t,s}\right)_{\lambda\Theta},\forall s\geq 0,\forall t\in\langle\langle s\rangle\rangle$$

$(P_{s,s}:=I_{k+1})$, by induction on t, $\forall s\geq 0$. Let $s\geq 0$.

$t=0$. Obvious.

$t=1$ (for $s\geq 1$). Obvious ($P_{s-1,s}=P_s$).

$t\longmapsto t+1$ ($t+1\leq s$). Suppose that

$$\min_{k\in S}\left(P_{s-t,s}\right)_{k\Theta}=\left(P_{s-t,s}\right)_{\lambda\Theta}.$$

We have

$$\left(P_{s-(t+1),s}\right)^{\{\Theta\}}=P_{s-t}\left(P_{s-t,s}\right)^{\{\Theta\}}=P_{s-t}\begin{pmatrix}\left(P_{s-t,s}\right)_{\lambda\Theta}\\ \left(P_{s-t,s}\right)_{a_{i_1}\Theta}\\ \left(P_{s-t,s}\right)_{a_{i_1}a_{i_2}\Theta}\\ \cdot\\ \cdot\\ \cdot\\ \left(P_{s-t,s}\right)_{a_{i_1}a_{i_2}\ldots a_{i_{k-1}}\Theta}\\ 1\end{pmatrix}=$$

$$=\begin{pmatrix}w_{i_1}^{(s-t)}\left(P_{s-t,s}\right)_{\lambda\Theta}+q_{i_1}^{(s-t)}\left(P_{s-t,s}\right)_{a_{i_1}\Theta}\\ w_{i_2}^{(s-t)}\left(P_{s-t,s}\right)_{\lambda\Theta}+q_{i_1}^{(s-t)}\left(P_{s-t,s}\right)_{a_{i_1}\Theta}+q_{i_2}^{(s-t)}\left(P_{s-t,s}\right)_{a_{i_1}a_{i_2}\Theta}\\ w_{i_3}^{(s-t)}\left(P_{s-t,s}\right)_{\lambda\Theta}+q_{i_1}^{(s-t)}\left(P_{s-t,s}\right)_{a_{i_1}\Theta}+q_{i_3}^{(s-t)}\left(P_{s-t,s}\right)_{a_{i_1}a_{i_2}a_{i_3}\Theta}\\ \cdot\\ \cdot\\ \cdot\\ w_{i_{k-1}}^{(s-t)}\left(P_{s-t,s}\right)_{\lambda\Theta}+q_{i_1}^{(s-t)}\left(P_{s-t,s}\right)_{a_{i_1}\Theta}+q_{i_{k-1}}^{(s-t)}\left(P_{s-t,s}\right)_{a_{i_1}a_{i_2}\ldots a_{i_{k-1}}\Theta}\\ w_{i_k}^{(s-t)}\left(P_{s-t,s}\right)_{\lambda\Theta}+q_{i_1}^{(s-t)}\left(P_{s-t,s}\right)_{a_{i_1}\Theta}+q_{i_k}^{(s-t)}\\ 1\end{pmatrix}$$

($\left(P_{s-(t+1),s}\right)^{\{\Theta\}}$ and $\left(P_{s-t,s}\right)^{\{\Theta\}}$ are the last columns of $P_{s-(t+1),s}$ and $P_{s-t,s}$, respectively). Now, it is easy to prove that

$$\left(P_{s-(t+1),s}\right)_{\lambda\Theta} \leq \left(P_{s-(t+1),s}\right)_{k\Theta}, \forall k \in S$$

$$\left(\left(P_{s-(t+1),s}\right)_{\lambda\Theta} = w_{i_1}^{(s-t)} \left(P_{s-t,s}\right)_{\lambda\Theta} + q_{i_1}^{(s-t)} \left(P_{s-t,s}\right)_{a_{i_1}\Theta}, \text{ etc.}\right).$$

Consequently,

$$\min_{k \in S} \left(P_{s-(t+1),s}\right)_{k\Theta} = \left(P_{s-(t+1),s}\right)_{\lambda\Theta}.$$

Second, from

$$\min_{k \in S} \left(P_{s-t,s}\right)_{k\Theta} = \left(P_{s-t,s}\right)_{\lambda\Theta}, \forall s \geq 0, \forall t \in \langle\langle s \rangle\rangle,$$

for $s = t = n$, we obtain

$$\min_{k \in S} \left(P_{0,n}\right)_{k\Theta} = \left(P_{0,n}\right)_{\lambda\Theta}.$$

To finish the proof, we use Theorem 1.1(i) and the fact that $\overline{\mu} = 1 - \mu$ and

$$P\left(X > n\right) = P\left(X_n \neq \Theta\right) = 1 - P\left(X_n = \Theta\right) = 1 - \left(P_{0,n}\right)_{\lambda\Theta}, \forall n \geq 0$$

(recall that the initial distribution is $\psi_0 = (1, 0, ..., 0)$). $\square$

Based on Theorem 4.9 all the results from A.1.1 on X from there, excepting the third inequality from Theorem 4.6, hold for X above as well – to obtain the result similar to the third inequality from Theorem 4.6, we must replace p^k with $q_{i_1} q_{i_2} ... q_{i_k}$, where $q_{i_s} := q_{i_s}^{(1)} = q_{i_s}^{(2)} = ..., \forall s \in \langle k \rangle$.

A.2.2. $X =$ the waiting time of pattern $\Theta = aba$ (here, the first symbol and the third one of Θ are identical) in an $a - b - c$ sequence of (independent) trials with

$$q_l^{(n)} = \text{the probability that } l \text{ occur in the } n\text{th trial}, \forall l \in \{a, b, c\}, \forall n \geq 1.$$

Following the Markov chain method, we associate X with the Markov chain $(X_n)_{n \geq 0}$ with the state space $S = \{\lambda, a, ab, \Theta\}$, initial distribution $\psi_0 = (1, 0, 0, 0)$, where $P(X_0 = \lambda) = 1$, and transition matrices (these are also lower Hessenberg matrices)

$$P_n = \begin{array}{c} \lambda \\ a \\ ab \\ \Theta \end{array} \begin{pmatrix} 1 - q_a^{(n)} & q_a^{(n)} & 0 & 0 \\ 1 - q_a^{(n)} - q_b^{(n)} & q_a^{(n)} & q_b^{(n)} & 0 \\ 1 - q_a^{(n)} & 0 & 0 & q_a^{(n)} \\ 0 & 0 & 0 & 1 \end{pmatrix}, n \geq 1.$$

It follows that

$$P\left(X > n\right) = P\left(X_n \neq \Theta\right), \forall n \geq 0.$$

Below we prove that the equations from Theorem 4.9 for X from there hold for X above as well. To see this, it is sufficient (see the proof of Theorem 4.9) to prove that

$$\min_{k\in S}\left(P_{s-t,s}\right)_{k\Theta}=\left(P_{s-t,s}\right)_{\lambda\Theta},\forall s\geq 0,\forall t\in\langle\langle s\rangle\rangle$$

($P_{s,s}:=I_4$). We show this by induction on t, $\forall s\geq 0$. Let $s\geq 0$.

$t=0$. Obvious.

$t=1$ (for $s\geq 1$). Obvious ($P_{s-1,s}=P_s$).

$t\longmapsto t+1$ ($t+1\leq s$). Suppose that

$$\min_{k\in S}\left(P_{s-t,s}\right)_{k\Theta}=\left(P_{s-t,s}\right)_{\lambda\Theta}.$$

We have

$$\left(P_{s-(t+1),s}\right)^{\{\Theta\}}=P_{s-t}\left(P_{s-t,s}\right)^{\{\Theta\}}=P_{s-t}\begin{pmatrix}\left(P_{s-t,s}\right)_{\lambda\Theta}\\ \left(P_{s-t,s}\right)_{a\Theta}\\ \left(P_{s-t,s}\right)_{ab\Theta}\\ 1\end{pmatrix}=$$

$$=\begin{pmatrix}\left(1-q_a^{(s-t)}\right)\left(P_{s-t,s}\right)_{\lambda\Theta}+q_a^{(s-t)}\left(P_{s-t,s}\right)_{a\Theta}\\ \left(1-q_a^{(s-t)}-q_b^{(s-t)}\right)\left(P_{s-t,s}\right)_{\lambda\Theta}+q_a^{(s-t)}\left(P_{s-t,s}\right)_{a\Theta}+q_b^{(s-t)}\left(P_{s-t,s}\right)_{ab\Theta}\\ \left(1-q_a^{(s-t)}\right)\left(P_{s-t,s}\right)_{\lambda\Theta}+q_a^{(s-t)}\\ 1\end{pmatrix}.$$

It follows that

$$\left(P_{s-(t+1),s}\right)_{\lambda\Theta}\leq\left(P_{s-(t+1),s}\right)_{k\Theta},\forall k\in S$$

$$\left(\left(P_{s-(t+1),s}\right)_{\lambda\Theta}=\left(1-q_a^{(s-t)}\right)\left(P_{s-t,s}\right)_{\lambda\Theta}+q_a^{(s-t)}\left(P_{s-t,s}\right)_{a\Theta},\text{ etc.}\right).$$

Therefore,

$$\min_{k\in S}\left(P_{s-(t+1),s}\right)_{k\Theta}=\left(P_{s-(t+1),s}\right)_{\lambda\Theta}.$$

Based on the above result, all the results from A.1.1 on X from there, excepting the third inequality from Theorem 4.6, hold for X above as well – to obtain the result similar to the third inequality from Theorem 4.6, we must replace p^k with $q_aq_bq_a$, where $q_l:=q_l^{(1)}=q_l^{(2)}=...$, $\forall l\in\{a,b,c\}$, and $\left\lfloor\frac{n}{k}\right\rfloor$ with $\left\lfloor\frac{n}{3}\right\rfloor$.

B. Results for $P(X=n)$.

B.1. Results based on monotone chains. We only consider this case (in the first example below, *i.e.*, B.1.1, we can work by induction instead of monotone matrices to prove Theorem 4.10). We give two examples. For these examples, B.1.1 and B.1.2, we consider that the possible outcomes for each trial are s and f. Set

$$p_n = \text{the probability that } s \text{ occur in the } n\text{th trial}$$

and

$$q_n = \text{the probability that } f \text{ occur in the } n\text{th trial}, \forall n \geq 1.$$

B.1.1. X = the waiting time of pattern $\Theta = s$ (X is a random variable with geometric distribution when trials are identically distributed).

Following the Markov chain method, we associate X with the Markov chain $(X_n)_{n\geq 0}$ with the state space $S = \langle\langle 1\rangle\rangle$ (we can also work with the state space $S = \{\lambda, \Theta\}$: λ instead of 0 and Θ instead of 1), initial distribution $\psi_0 = (1, 0)$, where $P(X_0 = 0) = 1$, and transition matrices

$$P_n = \begin{matrix} 0 \\ 1 \end{matrix} \begin{pmatrix} q_n & p_n \\ 0 & 1 \end{pmatrix}, n \geq 1.$$

It follows that

$$P(X = n) = P(X_0 < 1, X_1 < 1, ..., X_{n-1} < 1, X_n = 1), \forall n \geq 1.$$

Below we give other equations for $P(X = n)$.

THEOREM 4.10. *Under the above conditions we have*

$$P(X = n) = \bar{\alpha}(P_{0,n-1}Q_n) = \bar{\mu}(P_{0,n-1}Q_n) =$$
$$= \bar{\mu}\left((P_{0,n-1}Q_n)^{\{\overline{n}\}}\right) = 1 - (P_{0,n-1}Q_n)_{0\overline{n}} = (P_{0,n-1}Q_n)_{01}, \forall n \geq 1$$

(*see Section* 1 *for the fictive state* $\overline{n}$ *and for the matrix* Q_n).

Proof. Let $n \geq 1$. We have

$$Q_n = \begin{matrix} \\ 0 \\ 1 \end{matrix} \begin{matrix} \begin{matrix} 1 & \overline{n} \end{matrix} \\ \begin{pmatrix} p_n & q_n \\ 0 & 1 \end{pmatrix} \end{matrix}$$

(see before Theorem 1.2 again; see Theorem 1.5 and its proof again – we only used the fictive state $\overline{n}$).

We consider the chain $(U_m)_{m\geq 0}$ with the state spaces

$$S_m = \begin{cases} \langle\langle 1\rangle\rangle & \text{if } m \neq n, \\ \{1, \overline{n}\} & \text{if } m = n, \end{cases}$$

$m \geq 0$, initial distribution $\psi_0 = (1, 0)$, and transition matrices

$$R_m = \begin{cases} P_m & \text{if } m \neq n, n+1, \\ Q_n & \text{if } m = n, \\ C_{n+1} & \text{if } m = n+1, \end{cases}$$

$m \geq 1$, where

$$C_{n+1} := \begin{array}{c} \\ 1 \\ \overline{n} \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) \end{array}.$$

Since the chain $(U_m)_{m\geq 0}$ is monotone and

$$P(X=n) = P(X_0 < 1, X_1 < 1, ..., X_{n-1} < 1, X_n = 1) =$$
$$= P(U_0 < 1, U_1 < 1, ..., U_{n-1} < 1, U_n = 1) = P(U_n = 1) =$$
$$= 1 - P\left(U_n = \overline{n}\right),$$

the equations from Theorem 4.10 follow by Theorem 2.8. □

Remark 4.11. We have reasons similar to those from Remark 3.3 on the importance of equation $P(X=n) = \bar{\alpha}(P_{0,n-1}Q_n), \forall n \geq 1$, from Theorem 4.10.

THEOREM 4.12 (See also Theorem 2.5(i) in [17]). *Keeping the conditions from Theorem* 4.10 *we have*

$$E(X) = \sum_{n=1}^{\infty} n\bar{\alpha}(P_{0,n-1}Q_n)$$

$(P_{0,0} := I_2)$.

Proof. By Theorem 4.10,

$$E(X) = \sum_{n=1}^{\infty} nP(X=n) = \sum_{n=1}^{\infty} n\bar{\alpha}(P_{0,n-1}Q_n). \ \square$$

The equation from Theorem 4.12 is another bridge (see Theorem 4.3 and the paragraph after its proof again) between waiting time random variable theory and Markov chain theory. Both fields also can benefit by this connection. *E.g.*, if $E(X) < \infty$, then $n\bar{\alpha}(P_{0,n-1}Q_n) \to 0$ as $n \to \infty$ (equivalently, $n\bar{\alpha}(P_{0,n-1}Q_n) \nrightarrow 0$ as $n \to \infty$ implies $E(X) = \infty$). Hence, on Markov chain theory, we need results on Markov chains $(P_n)_{n\geq 1}$ with the property $n\bar{\alpha}(P_{0,n-1}Q_n) \to 0$ as $n \to \infty$ (or, equivalently, $n\bar{\alpha}(P_{0,n-1}Q_n) \nrightarrow 0$ as $n \to \infty$).

Let $m \geq 0$. Let $Y^{(m)}$ be another waiting time random variable which is similar to that from A.1.1, *i.e.*, the first trial of $Y^{(m)}$ and $(m+1)$th trial of X are identically distributed, the second trial of $Y^{(m)}$ and $(m+2)$th trial of X are also identically distributed, etc. We associate $Y^{(m)}$ with the Markov chain $(Z_n)_{n\geq m}$ with the state space $S = \langle\langle 1 \rangle\rangle$, initial distribution $\psi_0 = (1,0)$, where $P(Z_m = 0) = 1$, and transition matrices $V_n = P_n, \forall n > m$, where P_n, $n \geq 1$, are the transition matrices of chain $(X_n)_{n\geq 0}$ associated with X.

THEOREM 4.13. *Under the above conditions we have*

$$P(X = m+n) = P(X > m) P\left(Y^{(m)} = n\right), \forall m \geq 0, \forall n \geq 1.$$

In particular, if trials are identically distributed, we have

$$P(X = m+n) = P(X > m) P(X = n), \forall m \geq 0, \forall n \geq 1.$$

Proof. By Remark 4.4, Theorem 4.10, and A.1.2,

$$P(X = m+n) = \bar{\alpha}(P_{0,m+n-1}Q_{m+n}) = \bar{\alpha}(P_{0,m}) \bar{\alpha}(P_{m,m+n-1}Q_{m+n}) =$$

$$= P(X > m) P\left(Y^{(m)} = n\right), \forall m \geq 0, \forall n \geq 1.$$

If the trials are identically distributed, we have $P\left(Y^{(m)} = n\right) = P(X = n)$, $\forall m \geq 0, \forall n \geq 1$. □

Remark 4.14. (a) Theorem 4.13 gives one way to compute $P(X > m)$ easily when $m > 0$ (for $m = 0$, $P(X > 0) = 1$); indeed, for given $m, n \geq 1$,

$$P(X > m) = \frac{P(X = m+n)}{P\left(Y^{(m)} = n\right)} = \frac{q_1 q_2 ... q_{m+n-1} p_{m+n}}{q_{m+1} q_{m+2} ... q_{m+n-1} p_{m+n}} = q_1 q_2 ... q_m$$

and, in particular, if $p_1 = p_2 = ... := p$ and $q_1 = q_2 = ... := q$, *i.e.*, if the trials are identically distributed,

$$P(X > m) = \frac{P(X = m+n)}{P(X = n)} = \frac{q^{m+n-1}p}{q^{n-1}p} = q^m.$$

Further, since

$$P(X > m) = P(X = m+1) + P(X = m+2) + P(X = m+3) + ... =$$

$$= q_1 q_2 ... q_m p_{m+1} + q_1 q_2 ... q_{m+1} p_{m+2} + q_1 q_2 ... q_{m+2} p_{m+3} + ... =$$

$$= q_1 q_2 ... q_m (p_{m+1} + q_{m+1} p_{m+2} + q_{m+1} q_{m+2} p_{m+3} + ...),$$

we have

$$p_{m+1} + q_{m+1} p_{m+2} + q_{m+1} q_{m+2} p_{m+3} + ... = 1$$

(an equation which is useful – it is a generalization of $p + qp + q^2 p + ... = 1$ or, equivalently, of $1 + q + q^2 + ... = \frac{1}{1-q}$, where $p, q \geq 0$, $q \neq 1$, $p + q = 1$, and gives one way to obtain examples, as much as one likes, of convergent series) and, since

$$P(X > m) = 1 - P(X \leq m) =$$

$$= 1 - (P(X = 1) + P(X = 2) + P(X = 3) + ... + P(X = m)) =$$

$$= 1 - (p_1 + q_1 p_2 + q_1 q_2 p_3 + ... + q_1 q_2 ... q_{m-1} p_m),$$

we have

$$p_1 + q_1 p_2 + q_1 q_2 p_3 + ... + q_1 q_2 ... q_{m-1} p_m = 1 - q_1 q_2 ... q_m$$

(an equation which could be useful – it is a generalization of $p+qp+q^2p+...+q^{m-1}p=1-q^m$, where $p,q\geq 0$, $p+q=1$).

(b) By Theorem 4.13,

$$P\left(X=m+n\,|\,X>m\right)=P\left(Y^{(m)}=n\right),\forall m\geq 0,\forall n\geq 1,$$

and, in particular, if trials are identically distributed,

$$P\left(X=m+n\,|\,X>m\right)=P\left(X=n\right),\forall m\geq 0,\forall n\geq 1$$

(these are also lack-of-memory properties, see, for comparison, Theorem 4.7 and Definition 4.8).

B.1.2. $X=$ the waiting time of pattern $\Theta=ss$ (X is a random variable with geometric distribution of order 2 when the trials are identically distributed (see, *e.g.*, [1, p. 11])).

Following the Markov chain method, we associate X with the Markov chain $(X_n)_{n\geq 0}$ with the state space $S=\langle\langle 2\rangle\rangle$ (we also can work with the state space $S=\{\lambda,s,\Theta\}$: λ instead of 0, s instead of 1, and Θ instead of 2), initial distribution $\psi_0=(1,0,0)$, where $P(X_0=0)=1$, and transition matrices

$$P_n=\begin{array}{c}0\\1\\2\end{array}\begin{pmatrix}q_n & p_n & 0\\ 0 & q_n & p_n\\ 0 & 0 & 1\end{pmatrix},n\geq 1.$$

It follows that

$$P\left(X=n\right)=P\left(X_0<2,X_1<2,...,X_{n-1}<2,X_n=2\right),\forall n\geq 1.$$

Below we give other equations for $P(X=n)$.

THEOREM 4.15. *Under the above conditions we have, if* $p_n\geq q_n$ *(*$p_n\geq q_n\Longleftrightarrow p_n\geq\frac{1}{2}$*),* $\forall n\geq 1$,

$$P\left(X=n\right)=\overline{\alpha}\left(P_{0,n-1}Q_n\right)=\overline{\mu}\left(P_{0,n-1}Q_n\right)=$$
$$=\overline{\mu}\left(\left(P_{0,n-1}Q_n\right)^{\{\overline{n}\}}\right)=1-\left(P_{0,n-1}Q_n\right)_{0\overline{n}}=\left(P_{0,n-1}Q_n\right)_{02},\forall n\geq 2$$

$(P(X=0)=P(X=1)=0)$.

Proof. Let $n\geq 2$. We have (see before Theorem 1.2 again; see Theorem 1.5 and its proof again – we only need the fictive state $\overline{n}$)

$$Q_n=\begin{array}{c}\\0\\1\\2\end{array}\begin{array}{c}\begin{array}{cc}2 & \overline{n}\end{array}\\ \begin{pmatrix}0 & 1\\ p_n & q_n\\ 0 & 1\end{pmatrix}\end{array}.$$

The transition matrices of the above Markov chain are monotone, but Q_n is not. Nevertheless,

$$P_{n-1}Q_n = \begin{matrix} \\ 0 \\ 1 \\ 2 \end{matrix} \begin{matrix} \quad 2 \qquad\qquad \overline{n} \\ \begin{pmatrix} p_{n-1}p_n & 1-p_{n-1}p_n \\ q_{n-1}p_n & 1-q_{n-1}p_n \\ 0 & 1 \end{pmatrix} \end{matrix}$$

is monotone if $1-p_{n-1}p_n \leq 1-q_{n-1}p_n$, *i.e.*, since $p_n > 0$, if $p_{n-1} \geq q_{n-1}$.

We consider the chain $(U_m)_{m\geq 0}$ with state spaces

$$S_m = \begin{cases} \langle\langle 2\rangle\rangle & \text{if } m \neq n-1, \\ \{2, \overline{n}\} & \text{if } m = n-1, \end{cases}$$

$m \geq 0$, initial distribution $\psi_0 = (1, 0, 0)$, and transition matrices

$$R_m = \begin{cases} P_m & \text{if } m \neq n-1, n, \\ P_{n-1}Q_n & \text{if } m = n-1, \\ C_n & \text{if } m = n, \end{cases}$$

$m \geq 1$, where

$$C_n := \begin{matrix} \\ 2 \\ \overline{n} \end{matrix} \begin{matrix} 0 \quad 1 \quad 2 \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \end{matrix} .$$

Since the chain $(U_m)_{m\geq 0}$ is monotone if $p_n \geq q_n, \forall n \geq 1$, and

$$P(X = n) = P(X_0 < 2, X_1 < 2, ..., X_{n-1} < 2, X_n = 2) =$$

$$= P(U_0 < 2, U_1 < 2, ..., U_{n-2} < 2, U_{n-1} = 2) =$$

$$= P(U_{n-1} = 2) = 1 - P\left(U_{n-1} = \overline{n}\right),$$

the equations from Theorem 4.15 follow by Theorem 2.8. □

We can obtain a result related to Theorem 4.15. Indeed, keeping the conditions before Theorem 4.15, if we work with $P_{n-2}P_{n-1}Q_n$ instead of $P_{n-1}Q_n$ (see the above proof), then the chain $(R_m)_{m\geq 1}$ corresponding to this situation is monotone if $p_n \geq \frac{1}{3}$ ($p_n \geq \frac{1}{3} \Longleftrightarrow q_n \leq \frac{2}{3}$), $\forall n \geq 1$. Consequently, we have

$$P(X = n) = \bar{\alpha}(P_{0,n-1}Q_n) = \bar{\mu}(P_{0,n-1}Q_n) =$$

$$= \bar{\mu}\left((P_{0,n-1}Q_n)^{\{\overline{n}\}}\right) = 1 - (P_{0,n-1}Q_n)_{0\overline{n}} = (P_{0,n-1}Q_n)_{02}, \forall n \geq 3.$$

(Warning! These equations hold, here, for $n \geq 3$.) On the other hand, Theorem 4.15 leads to things similar to some from B.1.1, namely, Remark 4.11 and

Theorem 4.12, with only the difference that these refer to the case $p_n \geq q_n, \forall n \geq 1$. Moreover, Theorem 4.15 leads to a result related to Theorem 4.13; more precisely, keeping the same conditions as in Theorem 4.15, we have (for X from here, *i.e.*, from B.1.2)

$$P(X = m + n) \leq P(X > m) P\left(Y^{(m)} = n\right), \forall m \geq 0, \forall n \geq 1,$$

where the random variables $Y^{(m)}, m \geq 0$, are similar to those from A.1.1 and B.1.1, and, in particular, if the trials are identically distributed,

$$P(X = m + n) \leq P(X > m) P(X = n), \forall m \geq 0, \forall n \geq 1.$$

The general inequality leads, for $p_n \geq q_n, \forall n \geq 1$ (see Theorem 4.15), to

$$\frac{P(X = m + n)}{P\left(Y^{(m)} = n\right)} \leq P(X > m), \forall m \geq 0, \forall n \geq 1$$

with $P\left(Y^{(m)} = n\right) > 0$ (consequently,

$$\frac{P(X = m + n)}{P\left(Y^{(m)} = n\right)}$$

is a lower bound for $P(X > m)$, $\forall m \geq 0$, $\forall n \geq 1$ with $P\left(Y^{(m)} = n\right) > 0$ – note that $P\left(Y^{(m)} = 1\right) = 0$, $\forall m \geq 0$, and, since $p_n \geq q_n$ ($p_n \geq q_n \Longrightarrow p_n > 0$), $\forall n \geq 1$, we have $P\left(Y^{(m)} = n\right) > 0$, $\forall m \geq 0$, $\forall n \geq 2$) and to

$$P(X = m + n \mid X > m) \leq P\left(Y^{(m)} = n\right), \forall m \geq 0, \forall n \geq 1$$

(see also Remark 4.14 (b)). (The special inequality is left to the reader.)

For other examples of waiting time random variables, see, *e.g.*, [1] and [6] – our results, some of them with certain differences, could be extended to many other cases.

In [16–17], using our method based on ergodicity coefficients, etc., we gave upper bounds for certain probabilities and for the expectation of certain random variables (these studies could be used, in particular, to other ones, such as the study of expectation under the perturbation of probabilities). In this article, we used our method to obtain basic results for certain systems and random variables (some of results, such as Theorems 3.2 and 4.3, are interesting bridges between probability theory and Markov chain theory). This is the main contribution of our work.

REFERENCES

[1] N. Balakrishnan and M.V. Koutras, *Runs and Scans with Applications.* Wiley, New York, 2002.

[2] D.J. Daley, *Stochastically monotone Markov chains.* Z. Wahrscheinlichkeitstheorie Verw. Geb. **10** (1968), 305–317.

[3] R.L. Dobrushin, *Central limit theorem for nonstationary Markov chains,* I, II. Theory Probab. Appl. **1** (1956), 65–80, 329–383.

[4] P.I. Feder, *Problem solving: Markov chain method.* Industrial Engineering **6** (1974), 23–25.

[5] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd Edition. Wiley, New York, 1968.

[6] J.C. Fu and W.Y.W. Lou, *Distribution Theory of Runs and Patterns and Its Applications: a Finite Markov Chain Imbedding Approach.* World Scientific Publishing Co., River Edge, NJ, 2003.

[7] I.I. Glick, *A note on runs.* Z. Angew. Math. Mech. **61** (1981), 119–120.

[8] M. Iosifescu, *Finite Markov Processes and Their Applications.* Wiley, Chichester & Ed. Tehnică, Bucharest, 1980; corrected republication by Dover, Mineola, N.Y., 2007.

[9] D.L. Isaacson and R.W. Madsen, *Markov Chains: Theory and Applications.* Wiley, New York, 1976; republication by Krieger, 1985.

[10] J. Keilson, *Markov Chain Models – Rarity and Exponentiality.* Springer, New York, 1979.

[11] J. Keilson and A. Kester, *Monotone matrices and monotone Markov processes.* Stochastic Process. Appl. **5** (1977), 231–241.

[12] M.V. Koutras, *On a Markov chain approach for the study of reliability structures.* J. Appl. Probab. **33** (1996), 357–367.

[13] W. Kuo and M.J. Zuo, *Optimal Reliability Modeling: Principles and Applications.* Wiley, Hoboken, 2003.

[14] U. Păun, Δ-*ergodic theory and reliability theory.* Math. Rep. (Bucur.) **10(60)** (2008), 73–95.

[15] U. Păun, *A hybrid Metropolis-Hastings chain.* Rev. Roumaine Math. Pures Appl. **56** (2011), 207–228.

[16] U. Păun, $P(X_s \in A_s, X_{s+1} \in A_{s+1}, ..., X_t \in A_t)$ *in the Markov chain case: from an upper bound to a method.* Rev. Roumaine Math. Pures Appl. **57** (2012), 145–158.

[17] U. Păun, *Waiting time random variables: upper bounds.* Markov Process. Related Fields **19** (2013), 791–818.

Received 6 August 2015

Romanian Academy
"Gheorghe Mihoc – Caius Iacob" Institute,
of Mathematical Statistics
and Applied Mathematics,
Calea 13 Septembrie nr. 13,
050711 Bucharest 5, Romania
paun@csm.ro