Dedicated to Corresponding Member Ion Cuculescu
on the occasion of his 75^{th} birthday

$P(X_s \in A_s, X_{s+1} \in A_{s+1}, \ldots, X_t \in A_t)$ IN THE MARKOV CHAIN CASE: FROM AN UPPER BOUND TO A METHOD

UDREA PĂUN

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 $A_s, A_{s+1}, \ldots, A_t \subseteq S$, $A_s, A_{s+1}, \ldots, A_t \neq \emptyset, S$. We show that

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

where $\overline{\alpha} = 1 - \alpha$, α is the Dobrushin ergodicity coefficient, and $Q_{s,t} := Q_{s+1}Q_{s+2}\ldots Q_t$, $Q_{s+1}, Q_{s+2}, \ldots, Q_t$ are matrices which depend on (A_s, A_{s+1}) and P_{s+1}, (A_{s+1}, A_{s+2}) and $P_{s+2}, \ldots, (A_{t-1}, A_t)$ and P_t, respectively. This result and others (old or new results) lead to a new method for bounding certain probabilities $P(B)$, where, e.g., $B = \{X \in A\}$, X is a discrete random variable and $A \subseteq \mathbf{R}$, $0 < |A| < \infty$. To illustrate our method, we give upper bounds for the reliability of a k-out-of-v: F system and, more generally, of a weighted k-out-of-v: F system and for the reliability of a consecutive-k-out-of-v: F system.

AMS 2010 Subject Classification: 60J10, 60J20, 60E15, 65C60.

Key words: Markov chain, upper bound, ergodicity coefficient, Δ-ergodic theory, Markov chain method, distribution theory, reliability theory, k-out-of-v: F system, k-out-of-∞: F Markov chain, weighted k-out-of-v: F system, weighted k-out-of-∞: F Markov chain, consecutive-k-out-of-v: F system, consecutive-k-out-of-∞: F Markov chain.

1. AN UPPER BOUND

In this section, using ergodicity coefficients, we give an upper bound for $P(X_s \in A_s, X_{s+1} \in A_{s+1}, \ldots, X_t \in A_t)$, $0 \leq s < t$, in the finite Markov chain case, i.e., when $(X_n)_{n\geq 0}$ is a finite Markov chain with state space S and $A_s, A_{s+1}, \ldots, A_t \subseteq S$, $A_s, A_{s+1}, \ldots, A_t \neq \emptyset, S$. This result and others (old or new results) lead to a new method for bounding certain probabilities $P(B)$, where, e.g., $B = \{X \in A\}$, X is a discrete random variable and $A \subseteq \mathbf{R}$, $0 < |A| < \infty$ ($|A|$ denotes the cardinal of A), see Section 2.

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

Set

$$\langle m\rangle = \{1, 2, \ldots, m\}, \quad m \geq 1,$$

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

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

$$N_m = N_{m,m}, \quad \text{and} \quad S_m = S_{m,m}.$$

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 also shall refer to it as the (finite) Markov chain $(P_n)_{n\geq 1}$ (with state space $S = \langle r\rangle$). Define

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

(The entries of a matrix Z will be denoted Z_{ij}.)

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

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

(P_U, P^V, and P_U^V are matrices; e.g., if $m = 2$ and $n = 3$, then, e.g.,

$$P_{\{1\}} = \begin{pmatrix} P_{11} & P_{12} & P_{13} \end{pmatrix}, \quad P^{\{2\}} = \begin{pmatrix} P_{12} \\ P_{22} \end{pmatrix}, \quad \text{and} \quad P_{\{1\}}^{\{3\}} = (P_{13})),$$

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

($\alpha(P)$ is called the *Dobrushin ergodicity coefficient of* P (see, e.g., [2] or [7, p. 56])), and

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

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

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

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

Proof. (i) See, e.g., [7, p. 57] or [8, p. 144].

(ii) See, e.g., [2] or [8, p. 147].

(iii) See, e.g., [2], or [7, pp. 58–59], or [8, p. 145]. □

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

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

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

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

Below we give our the best result of this section.

THEOREM 1.2. *Under the above conditions we have*

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

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

Proof. Case 1. $\exists n \in \{s, s+1, \ldots, t\}$ such that $P(X_n \in A_n) = 0$. Obvious (because

$$P(X_s \in A_s, X_{s+1} \in A_{s+1}, \ldots, X_t \in A_t) \leq P(X_n \in A_n) = 0$$

and

$$\overline{\alpha}(P) \geq 0, \quad \forall P \in S_{g,h}).$$

Case 2. $P(X_n \in A_n) > 0$, $\forall n \in \{s, s+1, \ldots, t\}$. Let $p_u = ((p_u)_i)_{i \in S}$ be the probability distribution of the chain $(X_n)_{n \geq 0}$ at time u, $\forall u \geq 0$. Let $q_s = ((q_s)_i)_{i \in A_s \cup \{\overline{s}\}}$ be a probability vector (on $A_s \cup \{\overline{s}\}$), where

$$(q_s)_i := \begin{cases} (p_s)_i & \text{if } i \in A_s, \\ 1 - \sum\limits_{k \in A_s} (p_s)_k & \text{if } i = \overline{s}, \end{cases}$$

$\forall i \in A_s \cup \{\overline{s}\}$; we consider that $(q_s)_{\overline{s}}$ is the last component of q_s. Set

$$U_{t+1} = ((U_{t+1})_{ij})_{i \in A_t \cup \{\overline{t}\}, \, j \in S},$$
$$(U_{t+1})_{A_t} = (P_{t+1})_{A_t}, \quad (U_{t+1})_{\{\overline{t}\}} = (0, 0, \ldots, 0, 1);$$

we consider that $(U_{t+1})_{\{\overline{t}\}}$ is the last row of U_{t+1}.

Consider the Markov chain $(Y_n)_{n \geq s}$ with state spaces $A_s \cup \{\overline{s}\}, A_{s+1} \cup \{\overline{s+1}\}, \ldots, A_t \cup \{\overline{t}\}, S, S, \ldots$, initial probability distribution q_s, and transition matrices $(V_n)_{n\geq s+1}$, where

$$V_n := \begin{cases} Q_n & \text{if } n \in \{s+1, s+2, \ldots, t\}, \\ U_{t+1} & \text{if } n = t+1, \\ P_n & \text{if } n \geq t+2 \end{cases}$$

(this is a Markov chain with time varying state space, see, e.g., [7, p. 215]). We have

$$P(X_s \in A_s, X_{s+1} \in A_{s+1}, \ldots, X_t \in A_t) =$$

$$= \sum_{\substack{i_s \in A_s \\ i_{s+1} \in A_{s+1} \\ \vdots \\ i_t \in A_t}} P(X_s = i_s, X_{s+1} = i_{s+1}, \ldots, X_t = i_t) =$$

$$= \sum_{\substack{i_s \in A_s \\ i_{s+1} \in A_{s+1} \\ \vdots \\ i_t \in A_t}} P(X_s = i_s) P(X_{s+1} = i_{s+1} | X_s = i_s) \ldots P(X_t = i_t | X_{t-1} = i_{t-1}) =$$

$$= \sum_{\substack{i_s \in A_s \\ i_{s+1} \in A_{s+1} \\ \vdots \\ i_t \in A_t}} (p_s)_{i_s} P_{i_s, i_{s+1}} \ldots P_{i_{t-1}, i_t} = \sum_{\substack{i_s \in A_s \\ i_{s+1} \in A_{s+1} \\ \vdots \\ i_t \in A_t}} (q_s)_{i_s} Q_{i_s, i_{s+1}} \ldots Q_{i_{t-1}, i_t} =$$

$$= \sum_{\substack{i_s \in A_s \\ i_{s+1} \in A_{s+1} \\ \vdots \\ i_t \in A_t}} P(Y_s = i_s) P(Y_{s+1} = i_{s+1} | Y_s = i_s) \ldots P(Y_t = i_t | Y_{t-1} = i_{t-1}) =$$

$$= \sum_{\substack{i_s \in A_s \\ i_{s+1} \in A_{s+1} \\ \vdots \\ i_t \in A_t}} P(Y_s = i_s, Y_{s+1} = i_{s+1}, \ldots, Y_t = i_t) =$$

$$= P(Y_s \in A_s, Y_{s+1} \in A_{s+1}, \ldots, Y_t \in A_t) = P(Y_t \in A_t).$$

To finish the proof we show that

$$P(Y_t \in A_t) \leq \overline{\alpha}(Q_{s,t}).$$

To show this, let $\pi_s = (0, 0, \ldots, 0, 1) \in \mathbf{R}^{w_s}$, where $w_s := |A_s| + 1$ and $\pi_t = (0, 0, \ldots, 0, 1) \in \mathbf{R}^{w_t}$, where $w_t := |A_t| + 1$. Since $(Q_{s,t})_{\{\overline{s}\}} = \pi_t$, we have

$\pi_s Q_{s,t} = \pi_t$. Let q_u be the probability distribution of chain $(Y_n)_{n \geq s}$ at time $u - s$, $\forall u \geq s$. Further, by Theorem 1.1(ii), we have

$$\|q_t - \pi_t\|_1 = \|q_s Q_{s,t} - \pi_s Q_{s,t}\|_1 \leq \|q_s - \pi_s\|_1 \overline{\alpha}(Q_{s,t}) \leq 2\overline{\alpha}(Q_{s,t}).$$

On the other hand,

$$\|q_t - \pi_t\|_1 = \sum_{i \in A_t} (q_t)_i + 1 - (q_t)_{\overline{t}} = 2 \sum_{i \in A_t} (q_t)_i.$$

It follows that

$$\sum_{i \in A_t} (q_t)_i \leq \overline{\alpha}(Q_{s,t})$$

and, therefore,

$$P(Y_t \in A_t) = \sum_{i \in A_t} (q_t)_i \leq \overline{\alpha}(Q_{s,t}). \quad \square$$

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 of Theorem 1.2.

THEOREM 1.4. *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*

$$P(X_n < r) \leq \overline{\alpha}(P_{0,n}), \quad \forall n \geq 1$$

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

Proof. By the proof of Theorem 1.2,

$$P(X_0 < r, X_1 < r, \ldots, X_n < r) = P(X_n < r) \leq \overline{\alpha}(P_{0,n}), \quad \forall n \geq 1,$$

because, in this case, the chain $(Y_n)_{n \geq 0}$ from the proof of Theorem 1.2 ($s = 0$) is equal to $(X_n)_{n \geq 0}$ a.s. (almost surely) (in fact, we can work with $(X_n)_{n \geq 0}$ directly, i.e., without to use the intermediary chain $(Y_n)_{n \geq 0}$). $\square$

Remark 1.5. By the proof of Theorems 1.2 and 1.4,

$$P(X_n < r) = \frac{1}{2}\|p_n - \pi\|_1, \quad \forall n \geq 0,$$

where p_n is the probability distribution at time n of chain $(X_n)_{n \geq 0}$ from Theorem 1.4, $\forall n \geq 0$, and $\pi = (0, 0, \ldots, 0, 1) \in \mathbf{R}^r$.

2. A METHOD

In this section, based on the inequality $P(X_n < r) \leq \overline{\alpha}(P_{0,n})$, $\forall n \geq 0$, from Theorem 1.4, we give upper bounds for the reliability of a k-out-of-v: F system and, more generally, of a weighted k-out-of-v: F system and for the reliability of a consecutive-k-out-of-v: F system. To make this, we need upper bounds for $\overline{\alpha}(P_{0,v})$ (see Theorems 1.1 and 2.1); obviously, it is desirable to compute $\overline{\alpha}(P_{0,v})$, but this is impossible if k or v is too large.

Set

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

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

Set

$$(\{i\})_{i\in\{s_1,s_2,\ldots,s_t\}} = (\{s_1\},\{s_2\},\ldots,\{s_t\});$$
$$(\{i\})_{i\in\{s_1,s_2,\ldots,s_t\}} \in \mathrm{Par}(\{s_1,s_2,\ldots,s_t\}).$$

E.g., $(\{i\})_{i\in\langle m\rangle} = (\{1\},\{2\},\ldots,\{m\}) \in \mathrm{Par}(\langle m\rangle)$.

Below we give part of Theorem 1.8 from [15] and this in the stochastic case only; this is an important result and its proof is based on the G method from the Δ-ergodic theory. (For the general Δ-ergodic theory, see, e.g., [12–15] and the references therein.)

THEOREM 2.1. *Let* $P_1 \in S_{m_1,m_2}$, $P_2 \in S_{m_2,m_3}, \ldots, P_n \in S_{m_n,m_{n+1}}$. *Let* $\Delta_1 = (\langle m_1\rangle)$, $\Delta_2 \in \mathrm{Par}(\langle m_2\rangle), \ldots, \Delta_n \in \mathrm{Par}(\langle m_n\rangle)$, $\Delta_{n+1} = (\{i\})_{i\in\langle m_{n+1}\rangle}$. *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*

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

Then

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

(*Since* $L_1L_2\ldots L_n$ *is an* $1\times|\langle m_{n+1}\rangle|$ *matrix, it can be thought as being a row vector, but above we used and below we shall use the matrix notation for its entries instead of the vector one. Above the matrix notation* $(L_1L_2\ldots L_n)_{\langle m_1\rangle K}$ *was used instead of the vector one* $(L_1L_2\ldots L_n)_K$ *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]. $\square$

Set

$$\langle\langle m\rangle\rangle = \{0,1,\ldots,m\}, \quad m \geq 0,$$

and

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

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

Definition 2.2 (see also [12]). 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 $\operatorname{supp} \psi_0 \subseteq \langle\langle k-1 \rangle\rangle$, and transition matrices

$$P_n = \begin{pmatrix} p_n & \cdots & q_n & & & & & \\ & p_n & \cdots & q_n & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & p_n & & \cdots & q_n \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & & & p_n & q_n \\ & & & & & & & 1 \end{pmatrix}, \quad n \geq 1,$$

with $(P_n)_{i,i+w_n} = q_n$, $\forall n \geq 1$, $\forall i \in S$, $i + w_n \leq k$, 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, $1 \leq w_n \leq k$, $\forall n \geq 1$, 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-∞: F*. A 1-out-of-∞: 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., [9] and [10, p. 279]). Following the Markov chain method (see [3]; see, e.g., also [1, pp. 13–14], [5–6], and [9–10]), this system determines v stochastic matrices, say, $P_1, P_2, \ldots, P_v$ (we associate a stochastic matrix with each component of system), where P_n is identical with P_n from Definition 2.2, $\forall n \in \langle v \rangle$. To can work with Markov chains, since the matrices $P_1, P_2, \ldots, P_v$ do not determine a Markov chain, we consider a weighted k-out-of-∞: F Markov chain, say, $(X_n)_{n \geq 0}$ having the first v matrices even these ones ("∞" from weighted k-out-of-∞: F was suggested of 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 work, it follows that the reliability of above weighted k-out-of-v: F system, $R_v = R_v(k)$, is equal to $P(X_v < k)$, i.e.,

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

Consequently, to give upper bounds for $R_v = R_v(k)$, we can work in the weighted k-out-of-∞: F Markov chain framework. We shall use this framework and, therefore, we shall give upper bounds for $P(X_v < k)$.

Below we give a family of upper bounds for $P(X_v < k)$ of a k-out-of-∞: F Markov chain (equivalently, we give a family of upper bounds for reliability $R_v = R_v(k)$ of a k-out-of-v: F system (by definition, this is a v-component system which fails if and only if at least k of the v components fails, see, e.g., [10, p. 231])).

THEOREM 2.3. *Let* $(X_n)_{n\geq 0}$ *be a* k*-out-of-*∞*: F Markov chain. Let* $v \geq k$. *Let* $n_1, n_2, \ldots, n_k \geq 1$ (*they are integer numbers*) *such that* $v = n_1 + n_2 + \cdots + n_k$. *Then*

$$P(X_v < k) \leq 1 - (1 - p_1 p_2 \ldots p_{n_1})(1 - p_{n_1+1} p_{n_1+2} \ldots p_{n_1+n_2}) \ldots$$
$$\ldots (1 - p_{n_1+n_2+\cdots+n_{k-1}+1} p_{n_1+n_2+\cdots+n_{k-1}+2} \ldots p_{n_1+n_2+\cdots+n_{k-1}+n_k}).$$

Proof. By Theorem 1.4,

$$P(X_v < k) \leq \overline{\alpha}(P_{0,v}), \quad \forall v \geq 0.$$

We now give upper bounds for $\overline{\alpha}(P_{0,v})$ using Theorems 1.1(i) and 2.1. To make this, we consider the matrices

$$P_{0,n_1}, P_{n_1,n_1+n_2}, \ldots, P_{n_1+n_2+\cdots+n_{k-1},n_1+n_2+\cdots+n_{k-1}+n_k}$$

and partitions

$$\Delta_1 = (\langle\langle k\rangle\rangle), \quad \Delta_2 = (\{0\}, \langle k\rangle), \quad \Delta_3 = (\{0\}, \{1\}, \{2, 3, \ldots, k\}), \ldots,$$
$$\Delta_k = (\{0\}, \{1\}, \ldots, \{k-2\}, \{k-1, k\}), \quad \Delta_{k+1} = (\{i\})_{i\in\langle\langle k\rangle\rangle}$$

(recall that $(\{i\})_{i\in\langle\langle k\rangle\rangle} = (\{0\}, \{1\}, \ldots, \{k\})$). Since

$$P_{s,t} = \begin{pmatrix} p_{s+1}p_{s+2}\ldots p_t & & & & \\ & p_{s+1}p_{s+2}\ldots p_t & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & p_{s+1}p_{s+2}\ldots p_t & \\ & & & & 1 \end{pmatrix} + U_{(s,t)},$$

$\forall s,t,\ 0 \leq s < t$, where $U_{(s,t)}$ is a strictly upper triangular matrix, $\forall s,t$, $0 \leq s < t$, and (see Theorem 2.1; for labeling the rows and columns of matrices $L_1, L_2, \ldots, L_k$, we suppose that the partitions $\Delta_1, \Delta_2, \ldots, \Delta_{k+1}$ are ordered sets)

$$L_1 = L_1(P_{0,n_1}) = (0, 1 - p_1 p_2 \ldots p_{n_1})$$

($L_1 = L_1(P_{0,n_1})$ means that L_1 depends on P_{0,n_1}),

$$(L_2)_{\{\langle k\rangle\}} = (L_2(P_{n_1,n_1+n_2}))_{\{\langle k\rangle\}} = (0, 0, 1 - p_{n_1+1}p_{n_1+2}\ldots p_{n_1+n_2})$$

($(L_2)_{\{\langle k\rangle\}}$ is the last row of L_2 – this is the row $\langle k\rangle$ of L_2; we only need this row because $(L_1)_{\langle\langle k\rangle\rangle,\{0\}} = 0), \ldots,$

$$(L_k)_{\{\{k-1,k\}\}} = (L_k(P_{n_1+n_2+\cdots+n_{k-1},n_1+n_2+\cdots+n_{k-1}+n_k}))_{\{\{k-1,k\}\}} =$$
$$= (0, 0, \ldots, 0, 1 - p_{n_1+n_2+\cdots+n_{k-1}+1}p_{n_1+n_2+\cdots+n_{k-1}+2}\ldots p_{n_1+n_2+\cdots+n_{k-1}+n_k})$$

($(L_k)_{\{\{k-1,k\}\}}$ is the last row of L_k; we only need this row because, for $k \geq 2$, $(L_1 L_2 \ldots L_{k-1})^{\{\{0\},\{1\},\ldots,\{k-2\}\}}_{\{\langle\langle k\rangle\rangle\}} = (0, 0, \ldots, 0)$), we have

$$L_1 L_2 \ldots L_k = (0, 0, \ldots, 0, z),$$

where

$$z := (1 - p_1 p_2 \ldots p_{n_1})(1 - p_{n_1+1} p_{n_1+2} \ldots p_{n_1+n_2}) \ldots$$
$$\ldots (1 - p_{n_1+n_2+\cdots+n_{k-1}+1} p_{n_1+n_2+\cdots+n_{k-1}+2} \ldots p_{n_1+n_2+\cdots+n_{k-1}+n_k}).$$

Further, by Theorem 2.1,

$$\alpha(P_{0,v}) = \alpha(P_{0,n_1} P_{n_1,n_1+n_2} \ldots P_{n_1+n_2+\cdots+n_{k-1},n_1+n_2+\cdots+n_{k-1}+n_k}) \geq$$
$$\geq \sum_{J \in \Delta_{k+1}} (L_1 L_2 \ldots L_k)_{\langle\langle k\rangle\rangle J} = z.$$

Finally, by Theorem 1.1(i),

$$\overline{\alpha}(P_{0,v}) \leq 1 - z. \quad \square$$

Problem 2.4. How do we choose the numbers $n_1, n_2, \ldots, n_k$ in Theorem 2.3 to obtain an upper bound for $P(X_v < k)$ as small as possible? Obviously, we need that $1 - z$ (see the proof of Theorem 2.3 for the definition of z) be as small as possible; to fulfil this thing, one way is to choose the numbers $n_1, n_2, \ldots, n_k$ such that the products

$$p_1 p_2 \ldots p_{n_1},$$
$$p_{n_1+1} p_{n_1+2} \ldots p_{n_1+n_2}, \ldots,$$
$$p_{n_1+n_2+\cdots+n_{k-1}+1} p_{n_1+n_2+\cdots+n_{k-1}+2} \ldots p_{n_1+n_2+\cdots+n_{k-1}+n_k}$$

be as small as possible.

Example 2.5. Let $(X_n)_{n\geq 0}$ be a 10-out-of-∞: F Markov chain. If $p_1 = p_2 = \cdots = \frac{1}{2}$ (a homogeneous instance), then, for $v = 100$ and $n_1 = n_2 = \cdots = n_{10} = 10$ ($k = 10$), we have

$$P(X_{100} < 10) \leq 1 - \left(1 - \left(\tfrac{1}{2}\right)^{10}\right)^{10} \simeq 9.7228 \times 10^{-3}$$

while, for $v = 200$ and $n_1 = n_2 = \cdots = n_{10} = 20$, we have

$$P(X_{200} < 10) \leq 1 - \left(1 - \left(\tfrac{1}{2}\right)^{20}\right)^{10} \simeq 9.5367 \times 10^{-6}.$$

Further, if we consider $p_1 = p_2 = \cdots = p_{100} = \frac{1}{4}$ and $p_{101} = p_{102} = \cdots = \frac{1}{2}$ (a nonhomogeneous instance), then, for $v = 200$ and $n_1 = n_2 = \cdots = n_{10} = 20$, we have

$$P(X_{200} < 10) \leq 1 - \left(1 - \left(\tfrac{1}{4}\right)^{20}\right)^{5}\left(1 - \left(\tfrac{1}{2}\right)^{20}\right)^{5} \simeq 4.7684 \times 10^{-6}.$$

Using Theorem 2.1 directly, we can obtain better upper bounds for $P(X_v < 10)$ of each chain from Example 2.5. E.g., we consider, for $v = 100$, the first chain from Example 2.5. Taking, e.g., the matrices

$$P^{20},\ P^{20},\ P^{20},\ P^{20},\ P^{20}$$

and partitions

$$\Delta_1 = (\langle\langle 10\rangle\rangle), \quad \Delta_2 = (\{0,1\},\{2,3,\dots,10\}),$$
$$\Delta_3 = (\{0\},\{1\},\{2,3\},\{4,5,\dots,10\}),$$
$$\Delta_4 = (\{0\},\{1\},\{2\},\{3\},\{4,5\},\{6,7,\dots,10\}),$$
$$\Delta_5 = (\{0\},\{1\},\dots,\{5\},\{6,7\},\{8,9,10\}), \quad \Delta_6 = (\{i\})_{i\in\langle\langle 10\rangle\rangle},$$

then

$$P(X_{100} < 10) \le 1 - \left(1 - \frac{21}{2^{20}}\right)^5 \simeq 1.0013 \times 10^{-4}$$

because

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}^n = \begin{pmatrix} \frac{1}{2^n} & \frac{n}{2^n} \\ 0 & \frac{1}{2^n} \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} \frac{1}{2^n} & 1-\frac{1}{2^n} \\ 0 & 1 \end{pmatrix}, \quad \forall n \ge 1.$$

Moreover, since

$$\begin{pmatrix} p & q \\ 0 & p \end{pmatrix}^n = \begin{pmatrix} p^n & np^{n-1}q \\ 0 & p^n \end{pmatrix}, \quad \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} p^n & 1-p^n \\ 0 & 1 \end{pmatrix}, \quad \forall n \ge 1,$$

we can give other general upper bounds for $P(X_v < k)$ of a k-out-of-∞: F Markov chain in the homogeneous case. For the nonhomogeneous case, we can use a computer for evaluating the product of matrices

$$\begin{pmatrix} p_s & q_s \\ 0 & p_s \end{pmatrix} \begin{pmatrix} p_{s+1} & q_{s+1} \\ 0 & p_{s+1} \end{pmatrix} \cdots \begin{pmatrix} p_t & q_t \\ 0 & p_t \end{pmatrix}$$

for some $s \ge 1$ and $t > s$. Further, to obtain more and more general or special upper bounds for $P(X_v < k)$ of a k-out-of-∞: F Markov chain in the homogeneous case, we can consider

$$\begin{pmatrix} p & q & 0 \\ 0 & p & q \\ 0 & 0 & p \end{pmatrix}^n,$$

where $n \ge 1$. Etc.

To generalize Theorem 2.3 for weighted k-out-of-∞: F Markov chains we need the next remark.

Remark 2.6. Let $(P_n)_{n\ge 1}$ be a weighted k-out-of-∞: F Markov chain.

(a) By Theorems 1.1(iii) and 1.4,

$$P(X_v < k) \le \overline{\alpha}(P_{v_1,v_2}), \quad \forall v_1, v_2,\ 0 \le v_1 \le v_2 \le v$$

(if $v_1 = v_2 := h$, then we set $P_{h,h} = I_{k+1}$).

(b) If $P_{s+1}, P_{s+2}, \dots, P_t$ $(s < t)$ are matrices with weights $w_{s+1}, w_{s+2}, \dots, w_t$, respectively (see Definition 2.2 again), then, by induction,

$$(P_{s,t})_{i,i+j} = 0, \quad \forall i \in \langle\langle k-2 \rangle\rangle, \ \forall j, \ 0 < j < \min_{s<l\le t} w_l, \ i+j < k.$$

The result below is a generalization of Theorem 2.3 for $k, v \ge 2$.

THEOREM 2.7. *Let $(X_n)_{n\ge 0}$ be a weighted k-out-of-∞: F Markov chain. Let $k, v \ge 2$. Let $t \ge 1$. Let $0 \le m_0 < m_1 < \dots < m_t < m_{t+1} \le v$. Set $u_l = \min(w_{m_l+1}, w_{m_l+2}, \dots, w_{m_{l+1}})$, $\forall l \in \langle\langle t \rangle\rangle$. If*

$$u_0 + u_1 + \dots + u_{t-1} < k \le u_0 + u_1 + \dots + u_{t-1} + u_t$$

(*equivalently,*

$$u_0 + u_1 + \dots + u_{t-1} + 1 < |\langle\langle k \rangle\rangle| \le u_0 + u_1 + \dots + u_{t-1} + u_t + 1),$$

then

$$P(X_v < k) \le 1 - (1 - p_{m_0+1}p_{m_0+2}\dots p_{m_1})(1 - p_{m_1+1}p_{m_1+2}\dots p_{m_2})\dots$$
$$\dots(1 - p_{m_t+1}p_{m_t+2}\dots p_{m_{t+1}}).$$

Proof. It follows by Theorems 1.1(i), 1.4, and 2.1 and Remark 2.6. To see this, we take the matrices

$$P_{m_0,m_1}, P_{m_1,m_2}, \dots, P_{m_t,m_{t+1}}$$

and partitions

$$\Delta_1 = (\langle\langle k \rangle\rangle), \quad \Delta_2 = (\{0, 1, \dots, u_0 - 1\}, \{u_0, u_0 + 1, \dots, k\}),$$
$$\Delta_3 = (\{0\}, \{1\}, \dots, \{u_0 - 1\}, \{u_0, u_0 + 1, \dots, u_0 + u_1 - 1\},$$
$$\{u_0 + u_1, u_0 + u_1 + 1, \dots, k\}), \dots,$$
$$\Delta_{t+1} = (\{0\}, \{1\}, \dots, \{u_0 + u_1 + \dots + u_{t-2} - 1\},$$
$$\{u_0 + u_1 + \dots + u_{t-2}, u_0 + u_1 + \dots + u_{t-2} + 1, u_0 + u_1 + \dots + u_{t-2} + u_{t-1} - 1\},$$
$$\{u_0 + u_1 + \dots + u_{t-2} + u_{t-1}, u_0 + u_1 + \dots + u_{t-2} + u_{t-1} + 1, \dots, k\}),$$
$$\Delta_{t+2} = (\{i\})_{i \in \langle\langle k \rangle\rangle}.$$

(See the proof of Theorem 2.3 again.) □

Remark 2.8. If we have a k-out-of-v: F system, then $0 < k \le v$ while, if we have a weighted k-out-of-v: F system, then $0 < k \le v$ or $0 < v < k$ (see, e.g., [10, pp. 231 and 279]). Theorem 2.7 works when $2 \le k \le v$ or $2 \le v < k$.

Problem 2.9. How do we choose the numbers $m_0, m_1, \dots, m_{t+1}$ in Theorem 2.7 to obtain an upper bound for $P(X_v < k)$ as small as possible? (See Problem 2.4 again.)

We now deal with consecutive-k-out-of-v: F systems. Consider a system with v independent components which are linearly connected; such a system which malfunctions if and only if at least k consecutive components fail is called (*linear*) *consecutive-k-out-of-v: F system* (see, e.g., [10, p. 325]).

To give upper bounds for the reliability of a consecutive-k-out-of-v: F system (i.e., for the probability that this system work), we act as in the case of weighted k-out-of-v: F systems, i.e., since we can use the Markov chain method, we associate a consecutive-k-out-of-∞: F Markov chain with the consecutive-k-out-of-v: F system (as in the weighted k-out-of-v: F case, this association is not unique).

Definition 2.10 (see also [12]). A Markov chain with state space $S = \langle\langle k \rangle\rangle$ $(k \geq 1)$, initial probability distribution ψ_0 with $\operatorname{supp} \psi_0 \subseteq \langle\langle k-1 \rangle\rangle$, and transition matrices

$$P_n = \begin{pmatrix} p_n & q_n & & & & \\ p_n & & q_n & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & & \\ p_n & & & & & q_n \\ & & & & & 1 \end{pmatrix}, \quad n \geq 1,$$

is called *consecutive-k-out-of-∞: F*.

THEOREM 2.11. *Let $(X_n)_{n \geq 0}$ be a consecutive-k-out-of-∞: F Markov chain. Set $h = h(k, v) = [\frac{v}{k}]$, $\forall v \geq 0$. ($[x]$ denotes the integer part of x $(x \in \mathbf{R})$, i.e., $[x] := \max\{k \mid k \in \mathbf{Z}$ and $k \leq x\}$.) Then*

$$P(X_v < k) \leq (1 - q_1 q_2 \ldots q_k)(1 - q_{k+1} q_{k+2} \ldots q_{2k}) \ldots$$
$$\ldots (1 - q_{(h-1)k+1} q_{(h-1)k+2} \ldots q_{hk}), \quad \forall v \geq k.$$

Proof. By Theorem 2.1, taking the matrices

$$P_{m+1}, P_{m+2}, \ldots, P_{m+k} \ (m \geq 0)$$

and partitions

$$\Delta_1 = (\langle\langle k \rangle\rangle), \quad \Delta_2 = (\{0\}, \langle k \rangle), \quad \Delta_3 = (\{0\}, \{1\}, \{2, 3, \ldots, k\}), \ldots,$$
$$\Delta_k = (\{0\}, \{1\}, \ldots, \{k-2\}, \{k-1, k\}),$$
$$\Delta_{k+1} = (\{i\})_{i \in \langle\langle k \rangle\rangle},$$

we have

$$\alpha(P_{m,m+k}) \geq q_{m+1} q_{m+2} \ldots q_{m+k}.$$

In fact,

$$\alpha(P_{m,m+k}) = q_{m+1} q_{m+2} \ldots q_{m+k}, \quad \forall m \geq 0$$

(see [12, Theorem 2.13]); therefore, this is an example where Theorem 2.1 gives the best lower bound for $\alpha(P_{m,m+k})$. Further, by Theorems 1.1 ((i) and (iii)) and 1.4, we have

$$P(X_v < k) \le \overline{\alpha}(P_{0,v}) \le \overline{\alpha}(P_{0,k})\overline{\alpha}(P_{k,2k}) \ldots \overline{\alpha}(P_{(h-1)k,hk}) \le$$
$$\le (1 - q_1 q_2 \ldots q_k)(1 - q_{k+1} q_{k+2} \ldots q_{2k}) \ldots (1 - q_{(h-1)k+1} q_{(h-1)k+2} \ldots q_{hk}),$$

$\forall v \ge k$. $\square$

Remark 2.12. As to Theorem 2.11, if, in particular, $q_1 = q_2 \ldots := q$, then

$$P(X_v < k) \le (1 - q^k)^h, \quad \forall v \ge k$$

(by following a different approach, Muselli [11] also obtained this inequality).

To give better upper bounds for $P(X_v < k)$ of a consecutive-k-out-of-∞: F Markov chain $(P_n)_{n \ge 1}$ and of the others by Theorems 1.1, 1.4, and 2.1, we must obtain information about certain entries of the product $P_{m,n}$ for some $m \ge 0$ and $n > m$. Recall that it is possible that this information be obtained by computer (see, e.g., Example 2.5 and after it).

The Markov chain method is used for computing the probabilities of certain events which arise in the distribution theory and related fields (runs, patterns, reliability theory, hypothesis testing, quality control, etc., see, e.g., [1], [3], [4–6], and [9–11]). Theorems 1.1, 1.2, and 2.1, being general results, could be applied to give upper bounds for these probabilities (see Theorems 1.2 and 1.4 and this section again).

For other methods on lower and/or upper bounds for the probabilities of certain events which arise in the distribution theory and related fields, see, e.g., [1] and [10, Subsection 5.8] and the references therein. A drawback of the majority of these studies consists in the fact that they only refer to the homogeneous case. For our bounding method, it does not count if the chains are homogeneous or nonhomogeneous.

REFERENCES

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

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

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

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

[5] 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.

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

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

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

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

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

[11] M. Muselli, *On convergence properties of pocket algorithm.* IEEE Trans. Neural Networks **8** (1997), 623–629.

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

[13] U. Păun, G_{Δ_1,Δ_2} *in action.* Rev. Roumaine Math. Pures Appl. **55** (2010), 387–406.

[14] U. Păun, *General* Δ-*ergodic theory, with some results on simulated annealing.* Math. Rep. (Bucur.) **13(63)** (2011), 171–196.

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

Received 21 December 2011

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