PUBLICATIONS DE L'INSTITUT MATHÉMATIQUE
Nouvelle série, tome 28 (42), 1980, pp. 13–17

ALGORITHMICAL DEFINITION OF FINITE MARKOV SEQUENCE[1]

D. Banjević and Z. Ivković

0. Let $T = (t_1, t_2, \dots, t_N)$ be a finite binary sequence. Following von Mises ideas, A.N. Kolmogorov [1] defined the randomness of T with respect to the algorithm $R = (F, G, H)$ for selection of the subsequence of T. We give this definition in the following way:

The system of functions $F = (F_0, F_1, \dots F_{N-1})$, $F_0 = \text{const}$, defines a permutation $(x_1, x_2, \dots, x_N)$ of $(1, 2, \dots, N)$ which depends on T, by

$$x_i = F_{i-1}(x_1, t_{x1}; \dots, x_{i-1}, t_{xi-1}), \quad i = 1, 2, \dots, N.$$

The systems of functions $H = (H_0, H_1, \dots, H_N)$ and $G = (G_0, G_1, \dots, G_{N-1})$ have the properties: H_i, $G_i \in \{0, 1\}$, $H_0 = \text{const}$, $H_N = 1$, $H_i(x_0, t_{x1}; \dots; x_1, t_{xi}) \leq H_{i+1}(x_1, t_{x1}; \dots, x_{i+1}, t_{x_{i+1}})$, $G_0 = \text{const}$.

Let $s = s(T) = \min\{i : H_i = 1\}$ The system (F, G, H) defines the subset $A \subset \{1, 2, \dots, N\}$ in the following way: $x_k \in A$ if $1 \leq k \leq s$ and $G_{k-1}(x_1, t_{x1}, \dots x_{k-1}, t_{xk-1}) = 1$. Let $A = \{x_{i_1} x_{i_2}, \dots, x_{i_v}\}$, $x_{i_1} < x_{i_2} < \cdots x_{i_v}$. We select the subsequence $(t_{xi_1}, t_{xi_2}, \dots, t_{xi_v})$ of T by $R = (F, G, H)$.

The sequence T is (n, ε, p)-random $(1 \leq n \leq N, 0 < \varepsilon, 0 \leq p \leq 1)$ with respect to R if

$$v \geq n, \quad \text{and} \quad \left| \frac{1}{v} \sum_{k \in A} t_k - p \right| < \varepsilon$$

or if $v < n \cdot T$ is (n, ε, p)-random with respect to the system $\mathcal{R} = \{R_1, R_2, \dots\}$ if it is (n, ε, p)-random with respect to each $R_i \in \mathcal{R}$.

Another approach to the algoritmical definition of randomness was given in [2] and later developed for infinite set of sequnces ([3], [4]). However, for finite set of sequences which we consider here, this approach is too broad to be successful applied (see discution in [5]).

[1]This research was supported by Odsek za mat. i mat., meh. i astr. nauke, PMF Belgrade

Kolmogorov difinition of random sequence in [1] corresponds in a way to Bernoulli secuence $\mathcal{B}(p)$ in Probability theory. In this paper we define the randomness of T corresponding to homogenous Markov sequence $\mathcal{M}(\alpha,\beta)$ with the states $\{0,1\}$ and the transition matrix $\begin{bmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{bmatrix}$. In this definition we follow Kolmogorov's method.

1. It is reasonable that the definition of Markov sequence (MS) is based on the stability of the frequences of transition from 0 and 1. We select a subsequence $t_{i_1}, t_{i_2}, \dots, t_{i_k}$, $1 < i_1 < i_3 < \cdots i_k \le N$. Let $v_1 = \sum_{j-1}^{k} t_{i_j - 1}$ and $v_0 = k - v_1 = \sum_{j=1}^{k}(1 - t_{i_j - 1})$. Consider $\frac{1}{v_1}\sum_{j=1}^{k} t_{i_j-1} t_{i_j}$ -relative frequence of the transition from 1 to 1 and $\frac{1}{v_i}\sum_{j-\dot{u}}^{k}(1-t_{i_j-1})(1-t_{i_j})$-relative frequence of transition from 0 to 0. In accordance with the idea of MS, selection of a particular t_x in the subsequence should not depend on $t_x, t_{x+1}, \dots$. It means that the selection of t_i occurs before the selection of t_j for $i < j$.

Let $R = (G,H)$ be a system of functions $H = (H_0, H_1, \dots, H_N)$ and $G = (G_0, G_1 \cdots G_{N-1})$ with the properties H_i, $G_i \in \{0,1\}$, $H_0 = \text{const}$, $H_N = 1$, $H_i(t_1,\dots,t_i) \le H_{i+1}(t_1,\dots,t_{i+1})$, $G_0 = 0$.

Definition 1. The system $R = (G,H)$ is an algorithm for selection of the subsequence S of T, given by:

Let $s = s(T) = \min\{i : H_i = 1\}$. Let $A \subset \{1,2,\dots,N\}$ be defined by $j \in A$ iff $G_{i-1}(t_1,\dots,t_{i-1}) = 1$. Let $A = \{i_1, i_2, \dots, i - v\}$. Then $S = (t_{i_1}, t_{i_2}, \dots, t_{i_v})$.

By definition $2 \le i_1 < i_2 < \cdots i_v \le N$. From Definition 1 it follows that the algorithm $R = (G,H)$ is a particular Kolmogorov algorithm $R' = (F,G,H)$ where $F_{i-1} = i$, $i = 1,2,\dots,N$.

Definition 2. The secuence T is $(n_0, n_1, \varepsilon_0, \varepsilon_1, \alpha, \beta)$-Markov (denoted by $\mathcal{M}(n_0,n_1,\varepsilon_0,\varepsilon_1,\alpha,\beta)$, $(1 \le n_i \le N$, $0 < \varepsilon_i$, $i = 0,1$, $0 \le \alpha \le 1$, $0 \le \beta \le 1)$ with respect to R if

$$\text{(a)} \quad v_0 \ge n_0 \quad \text{and} \quad \Delta_0 = \left| \frac{1}{v_0} \sum_{j\in A} (1-t_{j-1})(1-t_j) - \alpha \right| < \varepsilon_0 \quad \text{or} \quad v_0 < n_0$$

$$\text{(b)} \quad v_1 \ge n_1 \quad \text{and} \quad \Delta_1 = \left| \frac{1}{v_1} \sum_{j\in A} t_{j-1} t_j - \beta \right| < \varepsilon_1 \quad \text{or} \quad v_1 < n_1$$

The sequence T is $\mathcal{M}(n_0,n_1,\varepsilon_0,\varepsilon_1,\alpha,\beta)$ with respect to the system $\mathcal{R} = \{R_1, R_3, \dots\}$ if it is $\mathcal{M}(n_0,n_1,\varepsilon_v,\varepsilon_1,\alpha,\beta)$ with respect to each $R_i \in \mathcal{R}$.

Definition 3. The sequence T is $(n_0,n_1,\varepsilon_0,\varepsilon_1,\alpha)$-Bernoulli (denoted by $\mathcal{B}(n_0,n_1,\ \varepsilon_0,\varepsilon_1,\alpha)$ with respect to $R(\mathcal{R})$ if it is $\mathcal{M}(n_0,n_1,\varepsilon_0,\varepsilon_1,\alpha,1-\alpha)$ with respect to $R(\mathcal{R})$.

Definition 3-follows from the idea that Bernoulli sequence is a particular Markov sequence for $\beta = 1 - \alpha$.

PROPOSITION 1. Let T be $\mathcal{B}(n_0, n_1, \varepsilon_0, \varepsilon_1, \alpha)$ with respect to $R = (G, H)$. Then T is $(\max\{n_0, n_1\}, \max\{\varepsilon_0, \varepsilon_1\}, 1-\alpha)$-random in the sence of Kolmogorov, with respect to the system $\{R_0, R_1\}$, $R_i = (F^j, G^j, H^j)$, $j = 0, 1$, where $H^j = H$, $j = 0, 1$, $F^j_{i-1} = i$, $i = 1, 2, \dots, N$, $j = 0, 1$ and

$$G^j_{i-1} = \begin{cases} G_{i-1} & t_{i-1} = j, \\ 0 & t_{i-1} = 1-j \end{cases}, \quad j = 0, 1.$$

PROOF. It is clear that R_0 is Kolmogorov algorithm. Let R select the subsequence $S = \{t_i\}$, $i \in A$ of T. Then R_0 selects the subsequence $S_0 = \{t_i\}$, $i \in B$ of S, which consits of elements proceeding zeros in T. Let S_0 have v elements. Evidently, $V = V_0$ and

$$\left|\frac{1}{v}\sum_{i\in B} t_i - (1-\alpha)\right| = \left|\frac{1}{v}\sum_{i\in B}(1-t_i) - \alpha\right| = \left|\frac{1}{v}\sum_{i\in A}(1-t_{i-1})(1-t_i) - \alpha\right| = \Delta_0.$$

Since T is $\mathcal{B}(n_0, n_1, \varepsilon_0, \varepsilon_1, \alpha)$ it follows that $v_0 < n_0$ or $v_0 \ge n_0$ and $\Delta_0 < \varepsilon_0$, e.i. $v < n_0$ or $v \ge n_0$ and $\left|\frac{1}{v}\sum_{i\in B} t_i - (1-\alpha)\right| < \varepsilon_0$. It means that T is $(n_0, \varepsilon_0, 1-\alpha)$-random with respect to R_0. Similary, T is $(n_1, \varepsilon_1, 1-\alpha)$-random with respect to R_1.

Generaly, let T be (n, ε, p)-random. Then T is (n, δ, p)- random for $m \ge n$, $\delta \ge \varepsilon$. Now since T is $(n_j, \varepsilon_j, 1-\alpha)$-random with respect to R_j, $j = 0, 1$, it means that T is $(\max\{n_0, n_1\}, \max\{\varepsilon_0, \varepsilon_1\}, 1-\alpha)$-random with respect to the system $\mathcal{R} = \{R_0, R_1\}$. Δ

2. In this section we consider the existence of at least one MS for a given system $\mathcal{R}$ with ρ algorithms.

Let $p(n, \varepsilon, \alpha) = P\left(\sup_{k\ge n}\left|\frac{S_k}{k} - \alpha\right| \ge \varepsilon\right)$ where random variable S_k have binomial distribution $b(k, \alpha)$.

PROPOSITION 2. Let system $\mathcal{R}$ have ρ algorithms. If

$$\rho < \frac{1}{p(n_0, \varepsilon_0, \alpha) + p(n_1, \varepsilon_1, \beta)}$$

then there exists at least one $\mathcal{M}(n_0, n_1, \varepsilon_0, \varepsilon_1, \alpha, \beta)$ sequence with respect to R.

PROOF. Consider Markov probability distribution on the set $\{T\}$, with given initial distribution and transition matrix $\begin{bmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{bmatrix}$. Let $P(R)(P(\mathcal{R}))$ be the probability that T is non-Markov with respect to $R(\mathcal{R})$. Then (using the same notation as in Definition 2)

$$P(R) = P((v_0 \ge n_0, \Delta_0 \ge \varepsilon_0) \cup (v_1 \ge n_1, \Delta_1 \ge \varepsilon_1)) \ge P(v_0 \ge n_0, \Delta_0 \ge \varepsilon_0) + \\ + P(v_1 \ge n_1, \Delta_1 \ge \varepsilon_1).$$

Let $\xi_1, \xi_2, \dots, \xi_i \in \{0,1\}$, be homogenous Markov chain with transition matrix $\begin{bmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{bmatrix}$. We select the sequence of indices $i_1, i_3 \dots, 1 \le i_1 < i_2 \dots$, such that $\xi_{i_1} = \xi_{i_2} = \dots = 0$ and that the selection of i_j is independant of ξ_k, $k > i_j$, $j = 1, 2, \dots$ Then the sequence $\xi_{i_1+1}, \xi_{i_2+1}, \dots$ is Bernoulli sequence where the probability of occuring 0 is α. Consider the sequence $\xi_1, \xi_2, \dots \xi_N$, as a part of infinite sequence $\xi_1, \xi_2, \dots$, and a subsequence $\xi_{i_1+1}, \xi_{i_2+1}, \dots \xi_{i_{v_0}+1}$ selected by R from $\xi_1, \xi_2, \dots \xi_N$. Let R^* be the algorithm defined for infinite sequence as the extension of R in the following way. R^* selects the same subsequence $\xi_{i_1+1}, \xi_{i_2+1}, \dots \xi_{i_k+1}$ as R until $k \le v_0$. For $k > v_0$ the selection is arbitrary (but in accordance with described rules of selection). Let $\eta_1, \eta_2, \dots$ be the selected subsequence. We define the stopping rule for R^* as

$$v_+^0 = n_0 \text{ if } \left| \frac{1}{n_0} \sum_{i-1}^{n_0} \eta_i - (1-\alpha) \right| \ge \varepsilon_0 \text{ and}$$

$$v_0^* = k, \ k > n_0, \text{ if } \left| \frac{1}{j} \sum_{i=0}^{j} \eta_i - (1-\alpha) \right| < \varepsilon_0, \ j = n_0, \ n_0+1, \dots k-1 \text{ and}$$

$$\left| \frac{1}{k} \sum_{i=1}^{k} \eta_i - (1-\alpha) \right| \ge \varepsilon_0.$$

Then

$$P(v_0^* \ge n_0, \Delta_0^* \ge \varepsilon_0) = P\left(\sup_{k \ge n^0} \left| \frac{S_k}{k} - (1-\alpha) \right| \ge \varepsilon_0 \right) =$$
$$p(n_0, \varepsilon_0, 1-\alpha) = p(n_0, \varepsilon_0, \alpha), \quad \left(\Delta_0^* = \left| \frac{1}{V_0^*} \sum_{i=1}^{v_0} \eta_i - (1-\alpha) \right| \right).$$

If T is non-Markov with respect to R, than each infinite sequence begining with T is non-Markov with respect to R^*. So $P(v_0 \ge n_0, \Delta_0 \ge \varepsilon_0) \le p(n_0, \varepsilon_0, \alpha)$. In the same way $P(v_1 \ge n_1, \Delta_1^* \ge \varepsilon_1) \le p(n_1, \varepsilon_1, \beta)$ i.e. $P(R) \le p(n_0, \varepsilon_0, \alpha) + p(n_1, \varepsilon_1, \beta)$ and $P(\mathcal{R}) \le \sum_{R \in \mathcal{R}} P(R) \le P[p(n_0, \varepsilon_0, \alpha) + p(n_1, \varepsilon_1, \beta)]$. If $\rho[p(n_0, \varepsilon_0, \alpha) + p(n_1, \varepsilon_1, \beta)] < 1$, then $P(\mathcal{R}) < 1$ and the probability measure of the set of Markov secuence is $1 - P(\mathcal{R}) < 0$, i.e. there exists at least one Markov sequence with respect to $\mathcal{R}$. Δ

Kolmogorov [1] gave the estimation $p(n, \varepsilon, \alpha) \le 2e^{-n\varepsilon^{2(1-\varepsilon)}}$. If

$$\rho < \frac{1}{2}[e^{-n_0 \varepsilon_0^2 (1-\varepsilon_0)} + e^{-n_1 \varepsilon_1^2 (1-\varepsilon_1)}]^{-1}$$

than for each system with ρ algorithms and each α and β there exists $\mathcal{M}(n_0, n_1, \varepsilon_0, \varepsilon_1, \alpha, \beta)$ sequence.

REFERENCES

[1] Kolmogorov, A.N. (1963). *On tables of random numbers*, Sankhya, ser. A. 25 369–376.

[2] Колмогоров, А.Н. (1965) *Три подхода к определению понятия "колчество информации"* Пробл. пер. информации, T I, No. 1. 3–7.

[3] Martin-Lof, (1966) *The definition of random sequences*, Inform. and Control, v. 9. 602–619.

[4] Левин, Л. А. (1973) *О понятии случайной последовательности*, Доклады АНСССР Т. 212, No. 3, 538–550.

[8] Banjević D. i Ivković Z. (1978) *Jedna definicija pravilnosti i slučajnosti zasnovana na idejama A.N. Kolmogoroova*, Matematički Vesnik, 2, (15), 39–47.