Probabilistic functions and statistical equivalence of binary shift registers with random Markov input

Sergey Yu. Melnikov, Konstantin E. Samouylov

Peoples' Friendship University of Russia (RUDN University), 6, Miklukho-Maklaya St., Moscow, 117198, Russia

Abstract

We consider a binary non-autonomous shift register with a sequence of random variables connected into a simple homogeneous stationary Markov chain at the input. The expression is obtained for the probability function in the output sequence in the form of a fractional rational function, the arguments of which are the transition probabilities of the Markov chain. An equivalence relation arising in the case of the identical equality of the probability functions of the registers is described. The results known earlier for the case when the input sequence is a sequence of independent random variables are generalized.

Keywords

shift register, automata with random input, probability function

1. Introduction

In a number of problems of recognition and identification of automata [1], the case is considered when the input of the automaton is a sequence of random variables. Recognition or identification of an automaton is carried out by analysis of statistics at the output of an automaton. It is known [2] that if the input of the automaton is a sequence of independent identically distributed random variables, then the distribution of symbols of the output sequence is described by a function on the Markov chain. Such a function can be specified by gluing states of the Markov chain [3]. In [4] the problem of synchronizing of finite automata, the input of which is a Bernoulli sequence, was studied. In [5] the problem of recognizing of the output function for three classes of automata was considered under the conditions when the input of the automaton is a sequence of independent identically distributed random variables. For the case when the automaton is a shift register and the input sequence is Bernoulli [6], an expression is obtained for the probability of a symbol in the output sequence. This probability polynomially depends on the parameter of the Bernoulli scheme, the degree of which does not exceed the size of the register. The coefficients of the polynomial are given by the sums of the values of the output function on some subsets of register states.

Here we study the case when the symbols of the input sequence of the binary shift register are connected into a simple homogeneous stationary Markov chain with two free parameters.

Workshop on information technology and scientific computing in the framework of the X International Conference Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems (ITTMM-2020), Moscow, Russian, April 13-17, 2020

melnikov@linfotech.ru (S. Yu. Melnikov); samuylov-ke@rudn.ru (K. E. Samouylov)
0000-0002-6368-9680 (K. E. Samouylov)

Our goal is to obtain an expression for the probability of a symbol in the output sequence and describe the equivalence relation on the set of output functions that provide the identity of the desired probabilities. We show that the desired probability is a fractional rational function of the parameters of the Markov chain, and the equivalence relation is given by the vectors of sums of the values of the output functions on some subsets of register states.

2. Definitions and statement of the problem

Let V_n be the space of n-dimensional binary vectors, F_n be the set of Boolean functions of n arguments, $n = 1, 2, \ldots$. For $f(x_1, x_2, \ldots, x_n) \in F_n$ by $A_f = (X = \{0,1\}, V_n, Y = \{0,1\}, h, f)$ we denote the Moore automaton with set of states V_n, the transition function h, determined by the rule $h((a_1, \ldots, a_n), x) = (a_2, \ldots, a_n, x)$, where $x, a_i \in \{0,1\}$, $i = 1, 2, \ldots, n$, and output function $f(x_1, x_2, \ldots, x_n)$. The automaton A_f is a shift register with a size of n.

If the Bernoulli sequence of binary random variables $x^{(i)}$, $i = 1, 2, \ldots$ with distribution $P\left(x^{(i)} = 1\right) = p$, $0 < p < 1$, used as the input of the automaton A_f, then the output sequence of random variables $f\left(x^{(i)}, x^{(i+1)}, \ldots, x^{(i+n-1)}\right)$, $i = n+1, n+2, \ldots$ is stationary and the probability $P\left\{f\left(x^{(i)}, x^{(i+1)}, \ldots, x^{(i+n-1)}\right) = 1\right\}$ of the symbol "1" in the output sequence is given by the polynomial [6]:

$$\Phi_f(p) = \sum_{j=0}^{n} s_j p^j (1-p)^{n-j}, \tag{1}$$

where

$$s_k = \sum_{(x_1, x_2, \ldots, x_n) : \sum x_i = k} f(x_1, x_2, \ldots, x_n), \quad k = 0, 1, \ldots, n. \tag{2}$$

Suppose that the automaton A_f, $f \in F_n$, $n = 1, 2, \ldots$, receives a sequence of binary random variables $x^{(i)}$, $i = 1, 2, \ldots$, connected in a simple homogeneous stationary Markov chain with the transition probability matrix

$$\begin{pmatrix} 1-\lambda & \lambda \\ \xi & 1-\xi \end{pmatrix}, \qquad 0 < \lambda, \quad \xi < 1. \tag{3}$$

The sequence of random variables $f\left(x^{(i)}, x^{(i+1)}, \ldots, x^{(i+n-1)}\right)$, $i = n+1, n+2, \ldots$ is stationary and it makes sense to talk about the probability $P\left\{f\left(x^{(i)}, x^{(i+1)}, \ldots, x^{(i+n-1)}\right) = 1\right\}$ of the symbol "1" in the output sequence.

The function $P_f(\lambda, \xi) = P\left\{f\left(x^{(i)}, x^{(i+1)}, \ldots, x^{(i+n-1)}\right) = 1\right\}$ will be called the probabilistic function of the automaton A_f with a Markov dependence at the input. Our task is to obtain an expression for $P_f(\lambda, \xi)$ and break F_n into classes with the same probabilistic functions.

3. Calculation of the probability function

We divide the set V_n of all n-dimensional binary vectors, $n \geqslant 2$, into four classes, depending on the values of the first and last coordinates. Let us denote:

$$\begin{aligned} V^{00} &= \{(0,\alpha_2,\alpha_3,\dots,\alpha_{n-1},0)\}, \quad & V^{01} &= \{(0,\alpha_2,\alpha_3,\dots,\alpha_{n-1},1)\}, \\ V^{10} &= \{(1,\alpha_2,\alpha_3,\dots,\alpha_{n-1},0)\}, \quad & V^{11} &= \{(1,\alpha_2,\alpha_3,\dots,\alpha_{n-1},1)\}, \end{aligned} \tag{4}$$

where $\alpha_i = 0, 1, i = 2, \dots, n-1$.

To each vector from V_n we associate its bigram marking $(\nu_{00}, \nu_{01}, \nu_{10}, \nu_{11})$, where $\nu_{\alpha\beta}$ is the number of bigrams (α, β) encountered in it. Let us put

$$\nu_{\alpha\beta} = \nu_{\alpha\beta}(\alpha_1,\alpha_2,\alpha_3,\dots,\alpha_n) = \sum_{k=1}^{n-1} \delta((\alpha\beta),(\alpha_k\alpha_{k+1})), \tag{5}$$

where δ is the Kronecker symbol, $\alpha, \beta = 0, 1$.

In each of the four classes, we single out the vectors characterized by the same markings $(\nu_{00}, \nu_{01}, \nu_{10}, \nu_{11})$. To do this, we introduce the following notation:

Let S_{ij}^{00} be the set of vectors from V^{00} with markings $(\nu_{00}, \nu_{01}, \nu_{10}, \nu_{11}) = (n-j-i-1, i, i, j-i)$. We denote I_1 as the set of possible pairs of indices (i, j). We will describe it.

For $n = 2l+2$, $l = 0, 1, 2, \dots,$

$$I_1 = \{(i,j)\} = \begin{cases} (0,0); \\ j = 1,\dots,l; \quad i = 1,\dots,j; \\ j = l+1,\dots,2l; \quad i = 1,\dots,2l-j+1; \end{cases} \tag{6}$$

for $n = 2l+1$, $l = 0, 1, 2, \dots,$

$$I_1 = \{(i,j)\} = \begin{cases} (0,0); \\ j = 1,\dots,l; \quad i = 1,\dots,j; \\ j = l+1,\dots,2l; \quad i = 1,\dots,2l-j+1. \end{cases} \tag{7}$$

Let S_{ij}^{01} be the set of vectors from V^{01} with markings $(\nu_{00}, \nu_{01}, \nu_{10}, \nu_{11}) = (n-j-i-1, i, i-1, j-i)$. We denote I_2 as the set of possible pairs of indices (i, j). We will describe it.

For $n = 2l+2$, $l = 0, 1, 2, \dots$

$$I_2 = \{(i,j)\} = \begin{cases} j = 1,\dots,l+1; \quad i = 1,\dots,j; \\ j = l+2,\dots,2l+1; \quad i = 1,\dots,2l-j+2; \end{cases} \tag{8}$$

for $n = 2l+1$, $l = 0, 1, 2, \dots$

$$I_2 = \{(i,j)\} = \begin{cases} j = 1,\dots,l; \quad i = 1,\dots,j; \\ j = l+1,\dots,2l; \quad i = 1,\dots,2l-j+1. \end{cases} \tag{9}$$

Let S_{ij}^{10} be the set of vectors from V^{10} with markings $(\nu_{00}, \nu_{01}, \nu_{10}, \nu_{11}) = (n-j-i-1, i-1, i, j-i)$. The set of possible pairs of indices (i, j) is the same as I_2.

Let S_{ij}^{11} be the set of vectors from V^{11} with markings $(\nu_{00}, \nu_{01}, \nu_{10}, \nu_{11}) = (n-j-i-1, i, i, j-i-1)$. The set of possible pairs of indices, which we denote by I_3, is obtained from I_1 by replacing j with $n-j$.

Note that the index j in the whole introduced notation is equal to the weight (sum of ones) of the vectors from the sets under consideration. The following lemma shows that the space of all n-dimensional binary vectors is represented as a union of the introduced sets, and these sets do not intersect.

Lemma. The following relations hold.

1. $S_{ij}^{\alpha\beta} \cap S_{kl}^{\gamma\delta} \neq \emptyset$ if and only if $\alpha = \gamma$, $\beta = \delta$, $i = k$, $j = l$.
2. $V^{\alpha\beta} = \bigcup_{(ij)} S_{ij}^{\alpha\beta}$, $\alpha, \beta = 0, 1$, depending on the value (α, β) the union is taken from among the sets I_1, I_2, I_3.
3. $\left|S_{ij}^{00}\right| = \binom{n-j-1}{i}\binom{j-1}{i-1}, \quad (i,j) \in I_1;$

   $\left|S_{ij}^{01}\right| = \left|S_{ij}^{10}\right| = \binom{n-j-1}{i-1}\binom{j-1}{i-1}, \quad (i,j) \in I_2;$

   $\left|S_{ij}^{11}\right| = \binom{n-j-1}{i-1}\binom{j-1}{i}, \quad (i,j) \in I_3.$

Hereinafter, we assume that

$$\binom{n}{-1} = \begin{cases} 1, & n = -1, \\ 0, & n \neq -1. \end{cases}$$

Proof. The first two points of the lemma follow from the construction of sets $S_{ij}^{\alpha\beta}$. The proof of the third one is based on counting the number of binary vectors of fixed weight with a given number of series of zeros and ones [7].

For $D \subset V^n$ we denote

$$\left\|f/_D\right\| = \sum_{(x_1,x_2,\ldots,x_n)\in D} f(x_1, x_2, \ldots, x_n). \tag{10}$$

Theorem 1. Let $x^{(i)}$, $i = 1, 2, \ldots$ be a stationary Markov chain with the states $\{0, 1\}$ and the transition probability matrix

$$\begin{pmatrix} 1-\lambda & \lambda \\ \xi & 1-\xi \end{pmatrix}. \tag{11}$$

The probabilistic function $P_f(\lambda, \xi)$ of the automaton A_f with a Markov dependence at the input has the form

$$P_f(\lambda,\xi) = \frac{1}{\lambda+\xi}\left\{\sum_{I_1}(1-\lambda)^{n-j-i-1}\lambda^i\xi^{i+1}(1-\xi)^{j-i}\left\|f/_{S_{ij}^{00}}\right\| + \right.$$
$$+\sum_{I_2}(1-\lambda)^{n-j-i}\lambda^i\xi^i(1-\xi)^{j-i}\left(\left\|f/_{S_{ij}^{01}}\right\| + \left\|f/_{S_{ij}^{10}}\right\|\right) +$$
$$\left. + \sum_{I_3}(1-\lambda)^{n-j-i}\lambda^{i+1}\xi^i(1-\xi)^{j-i-1}\left\|f/_{S_{ij}^{11}}\right\|\right\}. \tag{12}$$

To prove the Theorem 1, we use the lemma, the formulas for the total and conditional probability, and the well-known [8] form of the stationary distribution vector of the input sequence

$$\left(P\left(x^{(i)}=0\right), P\left(x^{(i)}=1\right)\right)=\left(\frac{\xi}{\lambda+\xi}, \frac{\lambda}{\lambda+\xi}\right). \tag{13}$$

4. The relation of statistical equivalence with Markov dependence at the input and its properties

By analogy with how this was done in [5], we introduce the equivalence relation on the set F_n. We call the functions f and g from F_n statistically equivalent for a Markov input dependence, having adopted the notation $f \overset{\Delta}{=} g$ for this case if the identity $P_f(\lambda,\xi)=P_g(\lambda,\xi)$ for $0<\lambda,\xi<1$ holds.

Obviously, the introduced relation is an equivalence relation breaking F_n into disjoint classes.

Vector $\overline{m}(f) = \left(\overline{m^{(1)}(f)}, \overline{m^{(2)}(f)}, \overline{m^{(3)}(f)}\right)$, where $\overline{m^{(i)}(f)} = \left(m_{ij}^{(k)}, (i,j)\in I_k\right)$, $k=1,2,3$, $m_{ij}^{(1)} = \left\| f/S_{ij}^{00} \right\|$, $(i,j)\in I_1$, $m_{ij}^{(2)} = \left\| f/S_{ij}^{01} \cup S_{ij}^{10} \right\|$, $(i,j)\in I_2$, $m_{ij}^{(3)} = \left\| f/S_{ij}^{11} \right\|$, $(i,j)\in I_3$, and the order of enumeration of the sets I_1, I_2, I_3 is fixed, for example, lexicographical, let's call the Markov weight structure of the Boolean function f.

Theorem 2. Two Boolean functions are $\overset{\Delta}{=}$-equivalent if and only if their Markov weight structures coincide.

Proof. Let us consider the system of real functions defined on the square $0<\lambda,\xi<1$:

$$\begin{cases} R_{ij}^{(1)}(\lambda,\xi)=\dfrac{1}{\lambda+\xi}(1-\lambda)^{n-j-i-1}\lambda^{i}\xi^{i+1}(1-\xi)^{j-i}, & (i,j)\in I_1 \\ R_{ij}^{(2)}(\lambda,\xi)=\dfrac{1}{\lambda+\xi}(1-\lambda)^{n-j-i}\lambda^{i}\xi^{i}(1-\xi)^{j-i}, & (i,j)\in I_2 \\ R_{ij}^{(3)}(\lambda,\xi)=\dfrac{1}{\lambda+\xi}(1-\lambda)^{n-j-i}\lambda^{i+1}\xi^{i}(1-\xi)^{j-i-1}, & (i,j)\in I_3 \end{cases}. \tag{14}$$

Denoting $\overline{R}(\lambda,\xi)=\left(R^{(1)},R^{(2)},R^{(3)}\right)$, where $R^{(k)}=\left(R_{ij}^{(k)}(\lambda,\xi),(i,j)\in I_k\right)$, $k=1,2,3$, we rewrite the expression (12) for the probability function in the form

$$P_f(\lambda,\xi)=\overline{R}(\lambda,\xi)\left(\overline{m}(f)\right)^T. \tag{15}$$

To complete the proof, it remains to note that the system (14) is a linearly independent system of functions on the square $0<\lambda,\xi<1$.

The proved theorem allows us to identify the $\overset{\Delta}{=}$ - equivalence class $[f]_{\overset{\Delta}{=}}$ with the vector $\overline{m}(f)$ of the Markov weight structure.

Since the coordinates of the vector $\overline{m}(f)$ are non-negative integers, it is not difficult to obtain the following description of the structure of the relation $\overset{\Delta}{=}$ - equivalence.

Theorem 3. The number of functions that are $\overset{\Delta}{=}$ - equivalent to a function f is determined by the expression

$$\left|[f]_{\overset{\Delta}{=}}\right| = \prod_{(i,j)\in I_1} \binom{\binom{n-j-1}{i}\binom{j-1}{i-1}}{\left\| f/S_{ij}^{00} \right\|} \prod_{(i,j)\in I_2} \binom{2\binom{n-j-1}{i-1}\binom{j-1}{i-1}}{\left\| f/S_{ij}^{01} \cup S_{ij}^{10} \right\|} \prod_{(i,j)\in I_3} \binom{\binom{n-j-1}{i-1}\binom{j-1}{i}}{\left\| f/S_{ij}^{11} \right\|}. \tag{16}$$

The number of $\overset{\Delta}{=}$ - equivalence classes is

$$\left|F_n/\overset{\Delta}{=}\right| = \prod_{(i,j)\in I_1} \left(\binom{n-j-1}{i}\binom{j-1}{i-1} + 1\right) \prod_{(i,j)\in I_2} \left(2\binom{n-j-1}{i-1}\binom{j-1}{i-1} + 1\right) \prod_{(i,j)\in I_3} \left(\binom{n-j-1}{i-1}\binom{j-1}{i} + 1\right) \tag{17}$$

For comparison the Table 1 presents the number of all Boolean functions, the number of $\overset{\Delta}{=}$-equivalence classes and the number of $\approx$-equivalence classes (equivalence at the Bernoulli input, see [5]) for $n = 2, 3, 4, 5$.

Table 1
The numbers of equivalence classes

n	2	3	4	5
$\lvert F_n \rvert$	16	256	65536	4294967296
$\left\lvert F_n/\overset{\Delta}{=} \right\rvert$	12	144	11664	8622400
$\left\lvert F_n/\approx \right\rvert$	12	64	700	17424

In view of the complexity of the expression for the number of $\overset{\Delta}{=}$-equivalence classes, it is of interest to estimate its growth at large n. Using the Stirling and Euler-Maclaurin formulas, we can obtain the following result.

Theorem 4. If $n \to \infty$, the relation

$$\left|F_n/\overset{\Delta}{=}\right| = \exp\left(\frac{5}{4}n^3 \ln n + O(n^3)\right). \tag{18}$$

5. Conclusion

The expression is obtained for the probabilistic function that describes the probability of a symbol in the output sequence of a binary shift register with a random binary variables connected into a simple homogeneous stationary Markov chain as input.

An equivalence relation is described on the set of output functions of binary shift registers that occurs when the corresponding probability functions are identically equal. The results obtained generalize those that were previously known for the case when the input sequence is a sequence of independent random variables.

Acknowledgments

The publication has been prepared with the support of the “RUDN University Program 5-100”. The reported study was funded by RFBR, project numbers 18-00-01555 (18-00-01685), 19-07-00933.

References

[1] S. Frenkel, Probabilistic model of control-flow altering based malicious attacks, in: O. Strichman, R. Tzoref-Brill (Eds.), Hardware and Software: Verification and Testing, Springer International Publishing, Cham, 2017, pp. 249–252.

[2] A. S. Davis, Markov chains as random input automata, The American Mathematical Monthly 68 (1961) 264–267.

[3] V. M. Zakharov, B. F. Eminov, S. V. Shalagin, Representation of markov’s chains functions over finite field based on stochastic matrix lumpability, in: 2016 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), 2016. doi:10.1109/ICIEAM.2016.7911662.

[4] V. V. Gusev, Synchronizing automata with random inputs, in: A. M. Shur, M. V. Volkov (Eds.), Developments in Language Theory, volume 8633, Springer International Publishing, Cham, 2014, pp. 68–75. doi:10.1007/978-3-319-09698-8_7.

[5] S. Y. Melnikov, K. E. Samouylov, The recognition of the output function of a finite automaton with random input, in: V. M. Vishnevskiy, D. V. Kozyrev (Eds.), Distributed Computer and Communication Networks, volume 919, Springer International Publishing, Cham, 2018, pp. 525–531. doi:10.1007/978-3-319-99447-5_45.

[6] B. A. Sevastyanov, The conditional distribution of the output of an automaton without memory for given characteristics of the input, Discrete Mathematics and Applications 4 (1994) 1–6. doi:10.1515/dma.1994.4.1.1.

[7] V. N. Sachkov, Combinatorial Methods in Discrete Mathematics, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1996.

[8] J. G. Kemeny, J. Snell, Finite Markov Chains, Springer-Verlag, 1976.