computing@computingonline.net
www.computingonline.net

Print ISSN 1727-6209
On-line ISSN 2312-5381
International Journal of Computing

GENERATION OF VALUES FROM DISCRETE PROBABILITY DISTRIBUTIONS WITH THE USE OF CHAOTIC MAPS

Marcin Lawnik, Arkadiusz Banasik, Adrian Kapczyński

Faculty of Applied Mathematics, Silesian University of Technology,
ul. Kaszubska 23, 44-100 Gliwice Poland,
marcin.lawnik@polsl.pl, arkadiusz.banasik@polsl.pl, adrian.kapczynski@polsl.pl, http://ms.polsl.pl

Paper history:
Received 6 September 2019
Received in revised form 14 February 2020
Accepted 21 February
Available online 31 March 2020

Keywords:
Piece-wise linear map;
Discrete random variable;
Bernoulli distribution.

Abstract: The values of random variables are commonly used in the field of artificial intelligence. The literature shows plenty of methods, which allows us to generate them, for example, inverse cumulative density function method. Some of the ways are based on chaotic projection. The chaotic methods of generating random variables are concerned with mainly continuous random variables. This article presents the method of generating values from discrete probability distributions with the use of properly constructed piece-wise linear chaotic map. This method is based on a properly constructed discrete dynamical system with chaotic behavior. Successive probability values cover the unit interval and the corresponding random variable values are assigned to the determined subintervals. In the next step, a piece-wise linear map on the subintervals is constructed. In the course of iterations of the chaotic map, consecutive values from a given discrete distribution are derived. The method is presented on the example of Bernoulli distribution. Furthermore, an analysis of the discussed example is conducted and shows that the presented method is the fastest of all analyzed methods.

1. INTRODUCTION

Many algorithms and calculating methods are based on random values. Unfortunately, the use of these methods is not always profitable, because of costs or time required for generating these numbers. In case of above-mentioned situations, we can apply pseudo-random values, which are using computers to generate them. The literature provides methods for continuous and discrete distributions. Algorithms and methods from the first group may be used in artificial intelligence algorithms, for example [1,2]. The other group of algorithms is used especially for data encryption, for example, stream ciphers [3,4]. The discrete values are often connected with binary values genration based on uniform distribution. In this case, the obtained values have to be of appropriate quality, which means they have to pass many statistical tests [5,6]. Also, discrete pseudo-random numbers may be used in other applications, for example, computer games, in which data processing form other probability distributions.

Many known methods use statistical relationships that allow you to generate pseudo-random numbers from discrete distributions, e.g. [7]. Some of the methods used for making pseudo-random values are based on chaotic mappings. The reason is connected with the properties of chaotic systems (such as behavior resembling the randomness) and the deterministic way in which values are obtained. Due to this reason, in the literature, one can find a lot of publications on the use of chaos in this type of issue. Most of such research works apply to continuous distributions [8,9,10] and generating strings of binary values [11,12,13]. Non-uniform discrete distributions are practically not considered in the literature with the use of chaotic mappings. For this reason, it seems reasonable to use chaotic maps and their properties to generate those types of values.

The main contribution of this work is the presentation of the method of generating pseudo-random values from a given discrete distribution. The method is using a properly constructed dynamic

system with chaotic behavior, which generates appropriate values as a result of iteration. The layout of the article consists of the following sections: Section 2 presents the mathematical foundations of chaotic mappings and discusses the research findings related to pseudo-random values which are generated using chaotic mappings; Section 3 presents a new method that allows to generate values from a given discrete distribution and an analysis with an example using the presented method; finally, conclusions are presented in section 4.

2. PRELIMINARIES AND RELATED WORKS

2.1 PRELIMINARIES

One of the most common and, at the same time, simplest dynamical system with chaotic behavior, is the asymmetric (skew) ten map. It is defined by the following formula (1).

$$x_{k+1} = f(x_k) = \begin{cases} \frac{x_k}{p}, & 0 < x_k \leq p \\ \frac{1-x_k}{1-p}, & p < x_k < 1 \end{cases}, \quad (1)$$

where $p \in (0,1)$. For any values of parameter p this dynamical system is chaotic. The Lyapunov exponent is calculated as:

$$\lambda = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |f'(x_i)| \quad (2)$$

for (1) is given by equation [14]:

$$\lambda = -p \ln p - (1-p)\ln(1-p). \quad (3)$$

Positive value of (2) is a necessary condition for dynamical system to appear chaos. The distribution (invariant density) of iterated variable of (1) is uniform, which is a direct result of the Frobenius-Perron equation [15].

A more general form of (1) was used in [16,17,18] for the construction of a compression algorithm, and expressed in the following way (4):

$$x_{k+1} = f(x_k) = \begin{cases} \frac{x_k}{p_1}, & x_k \in I_1 \\ \frac{x_k - p_1}{p_2}, & x_k \in I_2 \\ \vdots & \vdots \\ \frac{x_k - \sum_{i=1}^{n-1} p_i}{p_n}, & x_k \in I_n \end{cases}, \quad (4)$$

where $I_1 = [0, p_1]$, $I_i = \left[\sum_{j=1}^{i-1} p_j, \sum_{j=1}^{i} p_j\right]$, $\sum_{i=1}^{n} p_i = 1$.

Likewise (1), the distribution of the iterated variable of (4) is uniform, and the recurrence is chaotic with Lyapunov exponent equal to:

$$\lambda = -\sum_{i=1}^{n} p_i \ln p_i, \quad (5)$$

for any values of parameters p_i. Recurrence (4) was used as well in [19,20,21], for the purpose of construction of an encryption algorithm, which enabled the compression of the encrypted plaintext.

2.2 RELATED WORKS

The basic method of generating pseudo-random values from given probability distributions (both continuous and discrete) is the inverse cumulative distribution function method. This method requires a pseudo-random value $U \in [0,1]$ from a uniform distribution and by transforming it with a quantile function of a given probability distribution, one obtains the desired value. The whole can be described using the following equation [22]:

$$X = F^{-1}(U), \quad (6)$$

where $U \in [0,1]$ is the mentioned random variable value, F^{-1} is a quantile function and X is a random variable with distribution corresponding to F.

Method (6) in combination with mapping (1) was also applied in [8] for the derivation of chaotic recurrences, which made it possible to generate values from a given probability distribution. The recurrences are expressed as:

$$x_{k+1} = F^{-1}(T(F(x_k))), \quad (7)$$

where F is a cumulative distribution function of a given probability distribution, whereas T represents (1). Similar recurrences were analyzed in [9].

Furthermore, in [10] an algorithm for generating pseudo-random values based on the inversed method of "flattening" of probability distributions was shown. This method uses inverse transformations to chaotic recurrences.

The above mentioned juxtaposition of algorithms for generating pseudo-random values is concentrated on methods generating from continuous distributions. Literature also contains methods for generating values from discrete distributions. They focus mainly on receiving binary strings. The following is an example of such an algorithm [8, 9, 10]:

```
if x_k ≤ c then
    return 1;
else
    return 0;
end
```

Algorithm 1 – The method of generating binary values [11, 12, 13]

where x_k is the value obtained from a chaotic map, e.g., tent map (1), while c is fixed threshold value.

The above method is nothing else but a method of inverting a cumulative distribution function for the Bernoulli distribution, which is determined by means of the probability mass function [23]:

$$\rho(k) = \begin{cases} 1-p, & k=0 \\ p, & k=1 \end{cases} . \tag{8}$$

The method from the Algorithm 1 can be generalized to more than one threshold value c_i, thanks to which it will be possible to generate values from other discrete probability distributions.

3. PROPOSED METHOD

Let $\rho(x_i) = P(X = x_i) = p_i$ for $i = 1, 2, \ldots, n$ denote the probability mass function of discrete random variable X designated on a set W_X. Hence, the following equations must hold:

$$\sum_{i=1}^{n} p_i = 1 \text{ and } p_i \geq 0. \tag{9}$$

The values of p_i may be assigned to unit interval $[0, 1]$ as following:

$$I_1 = [0, p_1], I_i = \left[\sum_{j=1}^{i-1} p_j, \sum_{j=1}^{i} p_j \right] \text{ for } i = 2, 3, \ldots, n. \tag{10}$$

Accordingly, the recurrence in the form of (4) may be derived for the above mentioned separable unit interval. Next, the corresponding value x_i related to probability p_i may be assigned to each separable subinterval of the unit interval, i.e., each arm of considered map has assigned one value of X. Thus, in the course of the iterations of so constructed dynamical system, each of its consecutive iteration will derive a certain value of random variable x_i. In consequence, a sequence of the values assumed by random variable X are derived for the successive iterations of the map.

3.1 CORRECTNESS OF THE METHOD

The method is correct due to the properties of equation (4). Each subinterval on which (4) is defined is visited by the number of iterations that are proportional to its length in a given orbit. Because the invariant density of (4) is uniform, the successive lengths of the subintervals correspond to division (10).

Equation (4) is only one of the possible piece-wise linear maps, that enable the generation of values of the given discrete random variable X. Other forms of such maps are constructed in such manner, that their direction coefficient of each arm of the map is equal to $\frac{1}{p_i}$ or $-\frac{1}{p_i}$. Moreover, the order of the assignment probabilities to the unit interval in equation (10) may be changed. Thus, $2^n n!$ different linear maps may be derived, which, in the course of the iterations lead to the same distribution of their values (see Fig. 1).

Furthermore, for each of this $2^n n!$ piece-wise linear maps their Lyapunov exponent is also given by equation (5). Moreover, equation (5) expresses the entropy of the source that is a sequence of the values generated by means of described method (see [17]).

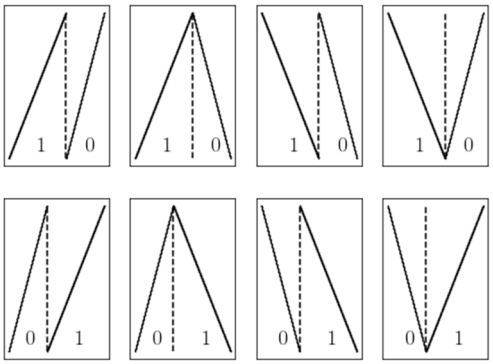


Figure 1 – Piece-wise linear maps which implement the generation of the discrete random variable from Bernoulli distribution

3.2 CASE STUDY

In accordance with the definition of the probability mass function for Bernoulli distribution (8), value 1 of random variable assumes probability p, whereas value 0 probability $1 - p$. Thus, recurrence in the form of (4) may be designated and the corresponding values 1 and 0 assigned to partition of unit interval (see Fig. 2). Next, by iterating the dynamic system described in the above mentioned manner, a binary sequence is obtained, generating values of random variables from Bernoulli distribution. Thus, Algorithm 1 is implemented by means of chaotic map (4).

The map plotted in Fig. 2 is not the only one of the chaotic maps that enables the procedure of generating values from Bernoulli distribution. Other recurrences that render the same result in the course of iterations are illustrated in Fig. 1, which are the same as used in [24] for the construction of one-time pad as a dynamical system. Accordingly, Algorithm 1 may be regarded as a particular case, and, in consequence, the chaotic maps in Fig. 1 as the generalization of the procedure of generating values from Bernoulli distribution.

Furthermore, in Fig. 3 the implementation of the described method by means of the recurrence from Fig. 2 and skew tent map (1) is plotted, starting at the same initial condition. The graph shows that the derived sequences of values are significantly different.

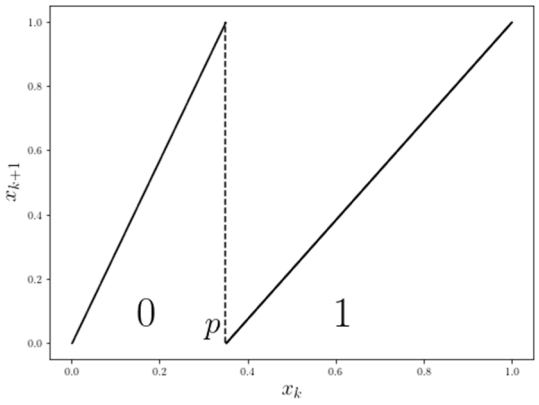


Figure 2 – Recurrence (4) with the values of 1 and 0 assigned to the successive subintervals

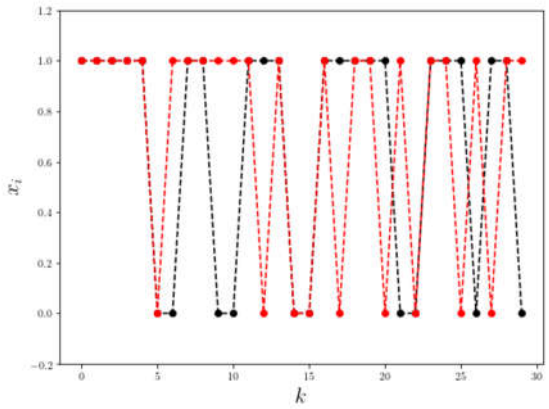


Figure 3 – Results of computed Bernoulli random variable values obtained by means of the function plotted in Fig. 2 (grey color) and equation (1) (black color) starting from the same initial condition

The presented method was compared with the following algorithms allowing to obtain a random variable from the Bernoulli distribution:

1. method 1: Algorithm 1 with tent map (1) with a threshold value $c = 0.7$
2. method 2: Algorithm 1 with mapping:

$$x_{k+1} = \begin{cases} \sqrt{2x_k}, & 0 < x_k \leq 0.5 \\ 1 - \sqrt{2 - 2x_k}, & 0.5 < x_k < 1 \end{cases} \quad (11)$$

with a threshold value $c = 0.7$. This mapping is similar to (1) - has an invariant density equal to 1 [14].

The above methods can be treated as special cases of the cumulative inversion method (6) with set chaotic maps (1) and (11). The results are presented in Table 1.

Table 1. Comparison of generation times 10^5 values from the Bernoulli distribution of the presented method with other algorithms using chaotic mappings. The value generation was repeated 100 times each time by selecting a random starting point value

Time	Method		
	Presented	1	2
Average (ms)	62	111	152
Std (ms)	1.34	8.46	7.01

It shows that the presented method is much faster than methods: 1 and 2. It is because a smaller number of comparisons are required for the presented method.

In turn, Figure 4 presents histograms of 0 and 1 values of generated binary strings. The obtained results confirm that all analyzed methods generate values from the Bernoulli distribution.

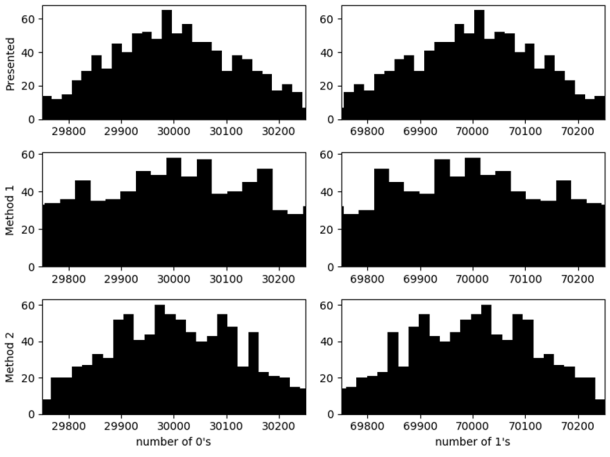


Figure 4 – Distributions of quantities of 0 and 1 for 10^3 binary strings with length 10^5 generated using the presented method and other mentioned algorithms

4. CONCLUSIONS

In this paper a method of generating values from discrete probability distributions with the use of piecewise linear chaotic maps is proposed. By

assigning the successive values of the discrete probability distribution to particular arms of the maps, the generated sequence of values from a given distribution is derived in the course of iterations. The method may be applied to any discrete distribution designated on the finite space of values. An example of generating values from Bernoulli distribution was presented. A comparative analysis was made. The presented method is the fastest of all mentioned in the article.

5. REFERENCES

[1] M. Woźniak, D. Połap, "Hybrid neuro-heuristic methodology for simulation and control of dynamic systems over time interval," *Neural Networks*, vol. 93, pp. 45-56, 2017. https://doi.org/10.1016/j.neunet.2017.04.013.

[2] R. Brociek, D. Słota, "Application and comparison of intelligent algorithms to solve the fractional heat conduction inverse problem," *Information Technology and Control*, vol. 45, issue 2, pp. 184-194, 2016.

[3] I. Gorbenko, A. Kuznetsov, Y. Gorbenko, S. Vdovenko, V. Tymchenko, M. Lutsenko, "Studies on statistical analysis and performance evaluation for some stream ciphers," *International Journal of Computing*, vol. 18, issue 1, pp. 82-88, 2019.

[4] W. Stallings, *Cryptography and Network Security: Principles and Practice*, 6th Edition, Pearson, 2013.

[5] L.E. Bassham, A.L. Rukhin, J. Soto, J.R. Nechvatal, M.E. Smid, E.B. Barker, S.D. Leigh, Mark Levenson, M. Vangel, D.L. Banks, N.A. Heckert, J.F. Dray, S. Vo, *SP 800-22 Rev. 1a. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*. Technical Report. National Institute of Standards & Technology, Gaithersburg, MD, USA, pp. 1-131, 2010.

[6] R.G. Brown, *Dieharder: A Random Number Test Suite*, 2020, [Online]. Available at: http://webhome.phy.duke.edu/~rgb/General/dieharder.php

[7] G. Marsaglia, W.W. Tsang & J. Wang, "Fast generation of discrete random variables," *Journal of Statistical Software*, vol. 11, issue 3, pp. 1-11, 2004.

[8] D. Lai, G. Chen, "Generating different statistical distributions by the chaotic skew tent map," *International Journal of Bifurcation and Chaos*, vol. 10, issue 6, pp. 1509-1512, 2000.

[9] M. Lawnik, "Generation of pseudo-random numbers from given probabilistic distribution with the use of chaotic maps," *Proceedings of the 6th International Eurasian Conference on Mathematical Sciences and Applications*, IECMSA-2017, Budapest, Hungary, August 15-18, 2017, pp. 1-6.

[10] M. Lawnik, "Generation of pseudo-random numbers with the use of inverse chaotic transformation", *Open Math.*, vol. 16, issue 1, pp. 16-22, 2018.

[11] A. Luca, A. Ilyas, A. Vlad, "Generating random binary sequences using tent map," *Proceedings of the IEEE International Symposium on Signals, Circuits and Systems (ISSCS 2011)*, Iasi, Romania, 2011, pp. 81–84.

[12] A. Ilyas, A. Vlad and A. Luca, "Statistical analysis of pseudorandom binary sequences generated by using tent map", *U. P. B. Sci. Bull.*, vol. 75, pp. 113-122, 2013.

[13] S. Šajić, N. Maletić, B.M. Todorović and M. Šunjevarić, "Random binary sequences in telecommunications," *Journal of Electrical Engineering*, vol. 64, issue 4, pp. 230-237, 2013.

[14] V.M. Anikin, S.S. Arkadaksky, S.N. Kuptsov, A.S. Remizov, L.P. Vasilenko, "Lyapunov exponent for chaotic 1D maps with uniform invariant distribution," *Bulletin of the Russian Academy of Sciences: Physics*, vol. 72, issue 12, pp.1684-1688, 2008.

[15] A. Lasota, M.C. Mackey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*, second ed., Springer, New York, 1993.

[16] M.B. Luca, A. Serbanescu, S. Azou, G. Burel, "A new compression method using a chaotic symbolic approach," *Proceedings of the IEEE-Communications*, Bucharest, Romania, 3-5 June 2004, pp. 1-6.

[17] N. Nagaraj, P.G. Vaidya, K.G. Bhat, "Arithmetic coding as a non-linear dynamical system," *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, issue 4, pp. 1013-1020, 2009.

[18] A. Pande, P. Mohapatra, J. Zambreno, "Using chaotic maps for encrypting image and video content," *Proceedings of the IEEE Int'l. Symp. Multimedia*, pp. 171-78, 2011.

[19] Y. Zhang, D. Xiao, H. Liu, H. Nan, "GLS coding based security solution to JPEG with the structure of aggregated compression and encryption," *Communications in Nonlinear Science and Numerical Simulation*, vol. 19, issue 5, pp. 1366-1374, 2014.

[20] K.W. Wong, Q. Lin, J. Chen, "Simultaneous arithmetic coding and encryption using chaotic maps," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 57, issue 2, pp. 146-150, 2010.

[21] Q. Lin, K.W. Wong, J. Chen, "Generalized arithmetic coding using discrete chaotic maps,"

International Journal of Bifurcation and Chaos, vol. 22, issue 10, pp. 1250256, 2012.

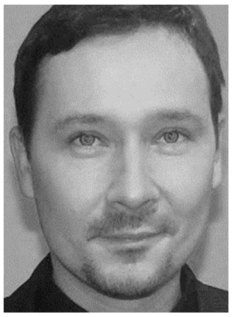
[22] L. Devroye, *Non-Uniform Random Variate Generation*, Springer-Verlag New York, 1986.

[23] C. Forbes, E. Merran, N. Hastings, B. Peacock, *Statistical Distributions*, fourth ed., John Wiley & Sons, New Jersey, 2011.

[24] N. Nagaray, "One-time pad as a nonlinear dynamical system," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, issue 11, pp. 4029-4036, 2012.

***Marcin Lawnik**, an Assistant Professor at the Faculty of Applied Mathematics, Silesian University of Technology; Ph.D. in Computer Science at the Częstochowa University of Technology (2015). Research interests: chaos based cryptography, chaos theory.*

***Arkadiusz Banasik**, a lecturer at the Faculty of Applied Mathematics, Silesian University of Technology. Research areas are concerned with artificial intelligence, especially in fuzzy logic, genetic algorithms, neural networks applied in IT field.*



***Adrian Kapczyński,** an Assistant Professor at the Faculty of Applied Mathematics, Silesian University of Technology. Ph.D. in Computer Science (2004, with distinction).*

Research interests: artificial intelligence, cloud computing, computer security, steganography and biometrics.