ISSN 1992-6502 (Print)
Vol. 18, no. 5 (66), pp. 54-56, 2014

Вестник УГАТУ
Vestnik UGATU

ISSN 2225-2789 (Online)
http://journal.ugatu.ac.ru

UDC 004.023+519.16

ON EFFICIENCY OF MODELING AN EQUIPROBABLY DISTRIBUTED SYSTEM OF RANDOM VARIABLES

EMIL YU. OREKHOV[1], **YURI V. OREKHOV**[2]

[1] emil.orekhov@bk.ru

Ufa State Aviation Technical University, Russia

Submitted 2014, June 10

Abstract. We suggest a modified algorithm of modeling an equiprobably distributed system of discrete random variables based on a non-equiprobably distributed system of discrete random variables with the same set of possible values. The efficiency of modeling is estimated.

Keywords: equiprobable generation; algorithm efficiency.

In [1, 2] a method is suggested which solves the problem of equiprobable generation of values for a system of discrete random variables. The solution is based on a non-equiprobable generator of another system of discrete random variables with the same set of possible values.

Namely, let $X=(X_1,...,X_n)$ be a system of n discrete random variables with a finite set of possible values $x^i=(x_1^i,...,x_n^i), i=1,...,N$. Let the probability distribution of the system be $P(X=x^i)=p_i, i=1,...,N, \sum_{i=1}^{N} p_i=1$.

Assume $p=\min_{i=1,...,N} p_i,\ p_i=p+\Delta_i,\ \Delta_i\ge 0, i=1,...,N$. Then, given the normalization

$$\sum_{i=1}^{N} p_i=1$$

we have

$$\sum_{i=1}^{N} \Delta_i=1-Np.$$

Consider a new system of discrete random variables $Z=(Z_1,...,Z_n)$, whose values are generated by the algorithm *EQPR*(*Z*).

Algorithm *EQPR*(*Z*)

1. Generate values for the system of random variables X according to its probability distribution. Let $X=x^i$ be the result of the step.
2. Generate values for an auxiliary system of random variables Y_i according to its probability distribution $P(Y_i=0)=\frac{\Delta_i}{p_i}, P(Y_i=1)=\frac{p}{p_i}$.
3. If $Y_i=1$, then set $Z=x^i$; else go to step 1 of the algorithm.

The system Z is proved to be equiprobably distributed, i.e.

$$P(Z=x^i)=\frac{1}{N}, i=1,...,N.$$

The chosen measure of efficiency of *EQPR*(*Z*) is the expectation $M[V]$ of the random variable V, which is the number of iterations required to deliver a realization of Z. It is proven that

$$M[V]=\frac{1}{Np}. \quad (1)$$

From (1) it follows that if p is rather small then delivering a realization of Z requires a lot of iterations. In this case the method efficiency is low. Therefore, increasing the efficiency of the suggested method of equiprobable generation is relevant.

In this paper we modify the described generator for an equiprobably distributed system of discrete random variables. We also show the modified algorithm can be substantially more efficient than the original *EQPR*(*Z*).

PROBLEM STATEMENT AND SOLUTION

Let $X=(X_1,...,X_n)$ be the system of discrete random variables defined above. Let A be a finite set of possible values of X, $|A|=N$. Let a set of equinumerous sets $A_k, |A_k|=L={N}/{K}, k=1,...,K$ be a partition of A. Let $p_j^k, k=1,...,K, j=1,...,L$ be a probability of the j-th element of A_k. Assume

$$p^k=\min_{j=1,...,L} p_j^k. \quad (2)$$

Let $W^k = (W_1^k, ..., W_n^k)$ be a system of discrete random variables with the set A_k of possible values and their corresponding probabilities

$$\frac{p_j^k}{\sum_{j=1}^{L} p_j^k}, \; k = 1, ..., K .$$

Consider a new system of discrete random variables $U = (U_1, ..., U_n)$ whose values are generated by the algorithm *EQPRM(U)*.

Algorithm *EQPRM(U)*

Generate equiprobably a number *k* of a subset of *A*.

Generate a realization w^k of the system W^k applying $EQPR(W^k)$ to the system W^k; note that the realization w^k coincides with a possible value x^i of the system X, $w^k = x^i$.

Assume $U = w^k = x^i$.

Obviously, the set of possible values of *U* is the set *A*.

Proposition 1

The system of random variables *U* is distributed equiprobably:

$$P(U = x^i) = \frac{1}{N}, \; i = 1, ..., N.$$

Proof. Let $x^i \in A^k$ and let α^k be an event of selecting number *k* on Step 1of the algorithm *EQPRM*(*U*) or, equally, selecting a subset A^k to be processed on the next step. Then

$$P(U = x^i) = P(U = x^i \,/\, \alpha^k) \cdot P(\alpha^k) .$$

According to *EQPRM*(*U*) we have

- $P(\alpha^k) = \frac{1}{K}$ by Step 1 of *EQPRM*(*U*);
- $P(U = x^i \,/\, \alpha^k) = \frac{1}{L}$ by Step 2 of *EQPRM*(*U*) taking into account W^k and the result of $EQPR(W^k)$.

Therefore

$$P(U = x^i) = \frac{1}{L} \cdot \frac{1}{K} = \frac{1}{LK} = \frac{1}{N},$$

since $L = \frac{N}{K}$.

The proposition is proved.

Assume V_M is the number of iterations required to obtain a realization of the system *U*. Specifically, V_M is the sum of the only iteration of Step 1 and all the iterations of $EQPR(W^k)$ on Step 2. Let the expectation $M[V_M]$ be a measure of efficiency of *EQPRM*(*U*).

According to (1), the expectation of the number of iterations of $EQPR(W^k)$ is

$$\frac{1}{L \cdot \frac{p^k}{\sum_{j=1}^{L} p_j^k}} .$$

Since it is the conditional expectation of the number of iterations on Step 2 of *EQPRM*(*U*) under the condition of equiprobable selection of A^k on Step 1, we obtain

$$M[V_M] = 1 + \sum_{k=1}^{K} \frac{1}{L \cdot \frac{p^k}{\sum_{j=1}^{L} p_j^k}} \cdot \frac{1}{K} =$$

$$= 1 + \frac{1}{KL} \sum_{k=1}^{K} \frac{\sum_{j=1}^{L} p_j^k}{p^k} = 1 + \frac{1}{N} \sum_{k=1}^{K} \frac{\sum_{j=1}^{L} p_j^k}{p^k}.$$

Therefore, the required measure of efficiency is

$$M[V_M] = 1 + \frac{1}{N} \sum_{k=1}^{K} \left(\frac{1}{p^k} \sum_{j=1}^{L} p_j^k \right). \qquad (3)$$

Regarding the upper and the lower bounds of $M[V_M]$ we claim that:

Proposition 2

$$2 \le M[V_M] \le 1 + \frac{1}{Np} . \qquad (4)$$

Proof. Obviously,

$$M[V_M] \ge 2 , \qquad (5)$$

since we always have at least Step1 of *EQPRM*(*U*) and at least one iteration on Step 2. The lower bound (5) is attainable since as $p_j^k = p^k, \; k = 1, ..., K, \; j = 1, ..., L$ we have

$$M[V_M] = 1 + \frac{1}{N} \sum_{k=1}^{K} \frac{1}{p^k} \sum_{j=1}^{L} p^k = 1 + \frac{K \cdot L}{N} = 1 + \frac{N}{N} = 2 .$$

Taking into account that $\frac{1}{p^k} \le \frac{1}{p}, \; k = 1, ..., K$, we can obtain the upper bound of $M[V_M]$ as follows.

$$M[V_M] = 1 + \frac{1}{N} \sum_{k=1}^{K} \left(\frac{1}{p^k} \sum_{j=1}^{L} p_j^k \right) \le 1 + \frac{1}{N} \cdot \frac{1}{p} \sum_{k=1}^{K} \sum_{j=1}^{L} p_j^k =$$

$$= 1 + \frac{1}{Np} \sum_{i=1}^{N} p_i = 1 + \frac{1}{Np},$$

i. e.

$$M[V_M] \le 1 + \frac{1}{Np} . \qquad (6)$$

Obviously, the upper bound (6) is attainable as $p^k = p, \; k = 1, ..., K$.

Combining (5) and (6), we get (4).

Now compare $M[V]$ and $M[V_M]$, which are the measures of efficiency of *EQPR*(*V*) and *EQPRM*(*U*), respectively, for the same system *X* of random variables. The minimum of $M[V]$ is 1, which is attainable, according to (1), as $p = \frac{1}{N}$. It follows that, obviously, since Step 1 of *EQPRM*(*U*) is always executed, then $M[V] < M[V_M]$ under "the best" and "the worst" conditions. However, from (1) and (3) we obtain the condition of *EQPRM*(*U*) being preferable to *EQPR*(*V*) according to the defined measure of efficiency: $M[V_M] < M[V]$ if

$$\sum_{k=1}^{K} \left(\frac{1}{p^k} \sum_{j=1}^{L} p_j^k \right) < \frac{1 - Np}{p}. \quad (7)$$

Example

Assume $N = 1000, K = 2$, $p = p^1 = 0.0001$, $p^2 = 0.001$, $\sum_{j=1}^{500} p_j^1 = 0.1, \sum_{j=1}^{500} p_j^2 = 0.9$.

Then, from (3) we get $M[V_M] = 2.9$, from (1) we get $M[V] = 10$, i.e. here *EQPRM* is 3.5 times more efficient than *EQPR*.

CONCLUSION

The suggested algorithm *EQPRM* can be substantially more efficient in practice than its predecessor *EQPR*. It follows from the proof of the Proposition 2 that, when partitioning a set *A* to blocks containing nearly equiprobable elements, we get the efficiency of *EQPRM* close to its lower bound, which equals 2.

REFERENCES

1. **E. Yu. Orekhov, Yu. V. Orekhov**, "The equiprobable generation of instances for the integer problem of scheduling jobs between unrelated parallel machines," in *Proc. of the 14th Int. Workshop on Computer Science and Information technologies CSIT'12* (Ufa – Hamburg – Norwegian Fjords, 2012), vol. 2, pp. 108–110, Ufa: UGATU, 2012.

2. **Орехов Э. Ю., Орехов Ю. В.** Об одном способе моделирования равновероятно распределенной системы дискретных случайных величин. // ITIDS'2013: Information Technologies for Intelligent Decision Making Support; Models and Algorithms of Applied Optimization: Proc. Int. Conf. and Intended Russian-German Workshop (May 21-25, Ufa, Russia). 2013. Т. 2. С. 63–65. **[** E. Yu. Orekhov, Yu. V. Orekhov, "On a method of modeling an equiprobably distributed system of random variables," (in Russian) in *Proc. of the Int. Conf. "Information Technologies for Intelligent Decision Making Support" and the Intended Russian-German Workshop "Models and Algorithms of Applied Optimization"*, May 21-25, Ufa, Russia, 2013, vol. 2, pp. 63-65. **]**

ABOUT AUTHORS

OREKHOV, Emil Yurievich, Docent of the Department of Computing Mathematics and Cybernetics. Dipl. Software Engineer (USATU, 1998), Cand. of Phys. & Math. Sci. (Samara State Aerospace University, 2002). Research in combinatorial optimization and heuristic algorithms.

OREKHOV, Yuri Vasilievich, Docent of the Department of Computing Mathematics and Cybernetics. Dipl. Physicist (Bashkir State University, 1973). Cand. of Tech. Sci. (Riga Polytechnical Institute, 1983). Research in combinatorial optimization and heuristic algorithms.

МЕТАДАННЫЕ

Название: Об эффективности моделирования равновероятно распределенной системы случайных величин

Авторы: Э. Ю. Орехов, Ю. В. Орехов

Организация: ФГБОУ ВПО «Уфимский государственный авиационный технический университет», Россия;

Email: emil.orekhov@bk.ru

Язык: английский.

Источник: Вестник УГАТУ. 2014. Т. 18, № 5 (66). С. 54–56. ISSN 2225-2789 (Online), ISSN 1992-6502 (Print).

Аннотация: Предложен модифицированный алгоритм моделирования равновероятно распределенной системы дискретных случайных величин, основанный на неравновероятно распределенной системе дискретных случайных величин с таким же набором возможных значений. Оценена эффективность моделирования.

Ключевые слова: равновероятная генерация; эффективность алгоритма.

Об авторах:

ОРЕХОВ Эмиль Юрьевич, доц. каф. выч. мат. и кибернетики. Дипл. инж.-программист (УГАТУ, 1998). Канд. физ.-мат. наук (Самарск. гос. аэрокосм. ун-т, 2002). Иссл. в области комбинаторной оптимизации и эвристических алгоритмов.

ОРЕХОВ Юрий Васильевич, доц. каф. выч. мат. и кибернетики. Дипл. физик (Баш. гос. ун-т, 1973). Канд. физ.-мат. наук (Рижский политехн. ин-т, 1983). Иссл. в области комбинаторной оптимизации и эвристических алгоритмов.