УДК 621.3.089, 621.317

# МЕТОДИКА ПОСТРОЕНИЯ ГЕНЕРАТОРА СЛУЧАЙНЫХ ЧИСЕЛ, РАСПРЕДЕЛЕННЫХ ПО ЗАКОНУ СИМПСОНА

Р.И. Годицкий, к.т.н. С.В. Рудаков
(представил д.ф.-м.н., проф. Е.Д. Прилепский)

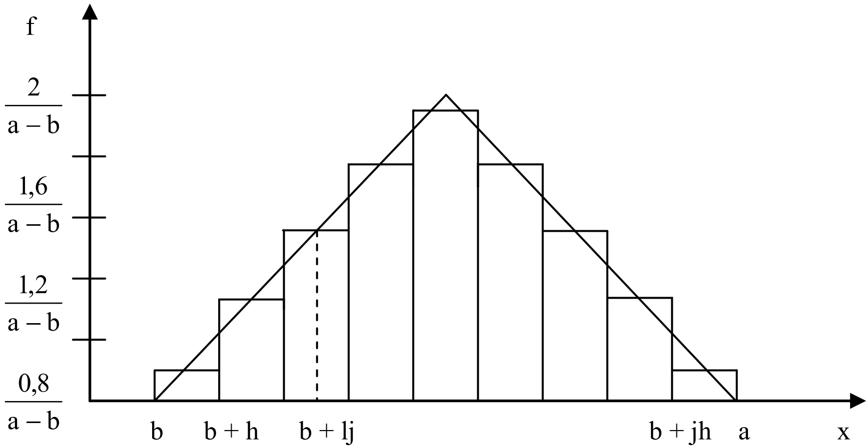
*Предлагается методика генерирования случайных чисел, распределенных по закону Симпсона, которая является наиболее эффективной, чем методика построения генератора случайных чисел на основе обратных функций.*

**Постановка проблемы**. При проведении имитационного моделирования для имитации влияния внешних воздействующих факторов, погрешностей средств измерительной техники, необходимы эффективные генераторы случайных чисел, распределенные по треугольному закону распределения. Генераторы, построенные методом обратных функций, показали низкую достоверность, поэтому возникла проблема создания нового генератора.

**Анализ литературы**. Для формирования случайных чисел на ПЭВМ, распределенных по закону Симпсона, наиболее часто применяют метод обратных функций, включающий генератор RND случайных чисел, распределенных по равномерному закону [1, 2]. Сформированный таким образом генератор, в области малых выборок чисел, менее 210, генерирует числа, которые как показывает практика, идентифицируются как распределенные по нормальному закону [3], а для выборок 240 – 250 чисел идентифицирует их как распределенные по закону Симпсона в 70 % случаев. Как известно [4], плотность вероятности закона Симпсона чисел, распределенных в интервале [a, b], имеет вид треугольника (рис. 1).

**Цель работы** – разработка алгоритма формирования случайных чисел, распределенных по закону Симпсона, и малые выборки которых идентифицируются как закон Симпсона.

**Методика построения генератора**. Поскольку общепринятые методики идентификации законов распределения случайных чисел основаны на анализе плотности вероятности, представленной в виде гистограмм, то построим алгоритм формирования гистограммы закона Симпсона. В идеальном варианте гистограмма должна быть такой, чтобы середины верхних отрезков интервалов гистограммы лежали на сторонах треугольника (рис. 1). Определим значения плотности вероятности на интервалах гистограммы. Учитывая, что площадь треугольника – вероятность появления всех чисел в интервале [a, b], т.е. численно равна единице, высота треугольника равна 2/(a – b).




$\frac{0,4}{a-b}$ Рис.1. Вид функции плотности вероятности по закону Симпсона

Тогда (рис.1), используя теорему о подобии треугольников [5], получаем выражение для плотности вероятности j-го интервала гистограммы

$$f_i = \frac{4\boldsymbol{l}_j}{(a-b)^2}, \tag{1}$$

где $\boldsymbol{l}_i$ – длина диапазона случайных чисел, заключенных между серединой j-го интервала гистограммы и началом выборки a; h – длина интервала, равная (a – b)/m; m – число интервалов; j – номер интервала, $j \le m/2$ для четного числа m, $j \le (m-1)/2$ для нечетного числа m.

Для четных m при $j > m/2$

$$f_j = f_{m-j}. \tag{2}$$

Для нечетных m

$$f_j = \frac{4}{(a-b)^2}\boldsymbol{l}_j \quad \text{при } j = 1, 2, \ldots, \frac{m-1}{2}; \tag{3}$$

$$f_{\frac{m+1}{2}} = \frac{2m-1}{(a-b)^2}h \quad \text{при} \quad j = \frac{m+1}{2}; \quad f_j = f_{m-j+1} \quad \text{при} \quad j = \frac{m+3}{2}, \ldots, m. \tag{4}$$

Длина диапазона $\boldsymbol{l}_j$ определяется как

$$\boldsymbol{l}_j = \frac{2j-1}{2}h. \tag{5}$$

С учетом (3) формула (1) преобразуется соответственно к виду

$$f_j = \frac{2(2j-1)}{(a-b)m}. \tag{6}$$

Алгоритм формирования случайных чисел, распределенных по закону Симпсона следующий. С учетом формулы (6) определим число $n_j$ чисел, попадающих в каждый интервал

$$n_j = n\frac{a-b}{m} f_j = \frac{2(2j-1)}{m^2} n. \tag{7}$$

Из формулы (7) следует, что при j = 1

$$n_1 = \frac{2n}{m^2}. \tag{8}$$

Таким образом, чтобы получить достоверную гистограмму опытного распределения Симпсона, необходимо, чтобы n и $n_1$ были такими, чтобы m было точно целым числом или близким к нему, т.е. должно выполняться соотношение

$$m \text{ (нечетное число)} \approx \sqrt{2n/n_1}, \text{ где } n_1=1;\ 2;\ 3;\ \ldots\ . \tag{9}$$

Если не удается получить идеальное соотношение, по формуле (9) определяется m как ближайшее целое нечетное число к результату извлечения корня квадратного из $2n/n_1$. Затем по формуле (7) определяется число $n_j$ чисел, попадающих в каждый нецентральный (округляется до ближайшего целого значения) интервал и с помощью генератора RND равномерных случайных чисел начинается заполнение интервалов числами, количество которых на каждом интервале ограничивается значением $n_j$. После заполнения всех интервалов числами, суммарное число которых является произведением целой части от деления n/m на число интервалов m, остаток чисел распределяется по интервалам. Заполнение интервалов остатком чисел осуществляется с помощью генератора равномерных случайных чисел при условии заполнения любого из интервалов не более одним числом.

При идентификации по методу наименьших квадратов построение теоретической функции по закону Симпсона осуществляется следующим образом. По выборке случайных n чисел организуется вариационный ряд, в котором $x_1$ = a, $x_n$ = b. Определяется математическое ожидание (или медиана) $\mu = \sum_{i=1}^{n} x_i \Big/ n$ и максимальные отклонения от среднего:

$$\Delta_{1max} = |\mu - a|; \quad \Delta_{2max} = |b - \mu|.$$

Выбирается наибольшее отклонение $\Delta_m$, как наибольшее значение из $\Delta_{1max}$ и $\Delta_{2max}$, и в качестве теоретической функции плотности вероятности строится функция в виде равнобедренного треугольника, одна вершина которого соответствует числу $\mu$, а две других числам $a_0 = \mu - \Delta_m$ и $b_0 = \mu + \Delta_m$ .

Расчет теоретической функции вероятности производится по формулам:

$$F = \frac{(x-\mu)^2}{2\Delta_m^2} + \frac{(x-\mu)}{\Delta_m} + \frac{1}{2},\ a_0 \le x \le \mu\ ;\ F = -\frac{(x-\mu)^2}{2\Delta_m^2} + \frac{(x-\mu)}{\Delta_m} + \frac{1}{2},\ \mu \le x \le b_0. \quad (10)$$

Идентификация производилась путем сравнения опытной функции $F_{оп}$ распределения с теоретическими (Симпсона, нормального закона, экспоненциального, равномерного, двумодального нормального, *arcsin*) по методу наименьших квадратов и наименьших максимальных отклонений.

$$\sum_{i-1}^{n}\left(F_{опi} - F_{теорi}(j)\right) = \min; \quad \sum_{i-1}^{n}\left(E_{опi} - E_{теорi}(j)\right) = \min, \qquad (11)$$

где j – порядковый номер теоретического закона распределения.

Согласующимся теоретическим законом распределения является тот, который даст наименьшее значение суммы отклонений по формуле (11).

**Выводы.** Предложена методика генерирования случайных чисел, распределенных по закону Симпсона, основанная на представлении теоретической функции плотности вероятности в виде ступенчатой функции и заполнении каждой ступеньки случайными числами, распределенными по равномерному закону. Предложенная методика позволяет получить выборки случайных чисел, которые идентифицируются как распределенные по закону Симпсона, начиная с n = 100, в то время как построение генератора по методу обратных функций позволяет распознавать закон Симпсона, начиная с n = 200, т.е. предложенная методика эффективнее методики на основе обратных функций.

## ЛИТЕРАТУРА

1. *Апанасович В.В., Коляда А.А., Чернявский А.Ф. Статистический анализ случайных потоков в физическом эксперименте. – Минск: Университетское, 1988. – 254 с.*
2. *Бусленко Н.П. Моделирование сложных систем. – М: Наука, 1978. – 400 с.*
3. *Соболь И.М. Численные методы Монте-Карло. – М: Наука, 1973. – 312 с.*
4. *Форсайт Дж., Малькогольм М., Моулер К. Машинные методы математических вычислений. – М: Мир, 1980. – 432 с.*
5. *Методика статистической обработки информации о надежности технических изделий на ЭЦВМ // ВНИИС ГКС СМ СССР, Издательство стандартов. – М., 1974. – 54 с.*

*Поступила 30.05.2003*

***ГОДИЦКИЙ Ремир Иосифович**, начальник учебной лаборатории ХВУ. В 1979 году окончил ХВВКУ. Область научных интересов – моделирование случайных процессов.*

***РУДАКОВ Сергей Валерьевич**, канд. техн. наук, преподаватель кафедры ХВУ. В 1996 году окончил ХВУ. Область научных интересов – моделирование случайных процессов.*