

**КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ НОРМАЛЬНО РАСПРЕДЕЛЕННЫХ
СЛУЧАЙНЫХ ВЕЛИЧИН****П. Н. Звягин, К. Н. Звягин (Санкт-Петербург)**

При решении ряда задач по имитационному моделированию важным элементом является компьютерное построение нормально распределенной случайной величины. В данной работе описано несколько подходов, позволяющих получить значения такой величины. Обращено внимание на особенности практического применения описываемых подходов, проведено сравнение программируемого метода полярных координат с методами получения нормально распределенной случайной величины пакета STATISTICA и из библиотеки Borland C++ 6.

Пусть требуется получить случайно распределенную случайную величину $N(\mu, \sigma)$, где μ и σ – параметры нормального распределения, соответствующие математическому ожиданию и среднеквадратическому отклонению.

Для решения этой задачи построим сначала две равномерно распределенные случайные величины, а затем преобразуем их в нормально распределенные с заданными параметрами.

Построение равномерно распределенных случайных величин

Для построения равномерно распределенных случайных величин используем так называемый конгруэнтный метод генерации последовательности случайных чисел. Этот метод относится к классу генераторов псевдослучайных чисел, т. е. таких чисел, последовательность которых можно восстановить при заданных начальных параметрах. Принцип генерирования линейной конгруэнтной последовательности был предложен Д.Х.Лемером в 1948 г., доработан Томсоном и Ротенбергом в 1958 и 1960 годах, соответственно [5].

Этот метод основан на использовании остатков от деления в качестве элементов случайной последовательности. Соотношение для получения равномерно распределенной случайной последовательности выглядит следующим образом:

$$X_{n+1} = (a \cdot X_n + c) \bmod m, \quad (1)$$

где \bmod обозначает операцию взятия остатка от деления;

X_n и X_{n+1} – два последовательных значения получаемой случайной величины,

$n \geq 0$;

a, c, m – параметры, правильный подбор которых обеспечивает увеличение периода последовательности случайных чисел и уменьшение их корреляции.

Это периодическая последовательность, поскольку свойство зацикливания присуще всем последовательностям, имеющим общий вид

$$X_{n+1} = f(X_n).$$

Существует теорема о выборе параметров в методе с целью получения некоторого известного периода.

Теорема. Длина периода линейной конгруэнтной последовательности равна m тогда и только тогда, когда:

1. c и m – взаимно простые числа;
2. $(a-1)$ кратно p для любого простого p , являющегося делителем m ;
3. $(a-1)$ кратно 4, если m кратно 4.

Доказательство этой теоремы приведено в [5].

В [4] предложены следующие соотношения для параметров:

$$X_0 \geq 0;$$

$$a = 2^l + 1, \text{ где } l \geq 2;$$

$$m = 2^p, \text{ } p \geq 2, \text{ и } m \text{ кратно } 4;$$

$$c > 0 \text{ и нечетно.}$$

При некоторых фиксированных параметрах a , c , m и изменении лишь стартовой точки X_0 получаются разные последовательности, являющиеся слабокоррелированными. При моделировании на компьютере в качестве стартовой точки возможно использование текущего значения секунд, полученного от системного таймера. Код на языке C++ для такого получения стартовой точки приведен ниже.

```
#include <ctime>
//----- time taking
    time_t * t1=NULL;
    time_t t2;
    tm * t3;

t2=time(t1);
t3=localtime(&t2);

// initializing start point x0 by time(seconds) meaning
x0=t3->tm_sec;
```

Полученная последовательность чисел будет распределена на интервале $[0, m]$.

Построение нормально распределенной величины с заданными параметрами μ , σ

В [2] указано, что следующие случайные величины независимы и распределены по нормальному закону с параметрами $\mu=0$, $\sigma=1$:

$$\eta_1 = \sin(2\pi\alpha_1) \cdot \sqrt{-2 \ln \alpha_2}, \quad (2)$$

$$\eta_2 = \cos(2\pi\alpha_1) \cdot \sqrt{-2 \ln \alpha_2}, \quad (3)$$

где α_1 , α_2 – две равномерно распределенные случайные величины. При этом величина α_1 распределена на интервале $[0, 1]$, а величина α_2 – на интервале $(0, 1]$.

Приведенный метод также называется методом полярных координат.

Получить величину α_1 можно описанным выше конгруэнтным методом, применив масштабирование делением полученных элементов последовательности на m . Таким образом будет осуществлен переход $[0, m] \rightarrow [0, 1]$.

Величину α_2 также можно получить конгруэнтным методом и для удовлетворения распределению на интервале $(0, 1]$ применить масштабирование со сдвигом:

$$X_n^{\alpha_2} = \frac{X_n + 1}{m + 1},$$

где $X_n^{\alpha_2}$ – n -ое значение случайной величины α_2 .

Из полученной величины Y , распределенной по нормальному закону $N(0,1)$ можно получить величину Z , распределенную по нормальному закону $N(b,|a|)$ с заданными параметрами b и a , путем следующего преобразования [3]:

$$Z = a \cdot Y + b.$$

При этом b можно интерпретировать как заданное математическое ожидание μ , а $|a|$ – как заданное среднеквадратическое отклонение σ .

Оценка полученных распределений. Сравнение с распределениями, полученными в пакете STATISTICA и при использовании функции Borland C++ 6

Последовательность операций, использующихся для получения нормально распределенной случайной переменной в пакете STATISTICA [1], выглядит следующим образом:

$$\mathbf{Randvar} = \mathbf{rnd}(1) \tag{4}$$

$$\mathbf{Normvar} = \mathbf{Vnormal}(\mathbf{Randvar}; m; \sigma^2).$$

Здесь **Randvar** – переменная, значения которой равномерно распределены на отрезке $[0,1]$, а **Normvar** – построенная на основе **Randvar** переменная, значения которой имеют нормальное распределение с параметрами m и σ^2 .

В Borland C++ 6, в одной из библиотек, существует функция, позволяющая получить значения нормально распределенной случайной величины.

Приведем пример ее использования:

```
#include <math.hpp>
z = RandG(0,1);
```

(5)

Здесь **RandG**(μ, σ) – функция для получения значений нормально распределенной случайной величины с параметрами μ, σ, z – может быть типа **double**.

Следует отметить, что во многих других версиях языка, например в стандартном выпуске Microsoft Visual C++ 6, такой функции нет.

Оценим полученные выборки на соответствие нормальному закону распределения по критерию χ^2 . Для теста построим последовательности псевдослучайных чисел по соотношению (2), также воспользуемся генератором нормально распределенных случайных величин пакета STATISTICA (4) и функцией **RandG** (5) из Borland C++ 6.

При использовании в качестве стартовых точек текущих значений секунд для генерации равномерно распределенных случайных последовательностей (1) получаем достаточно большое число вариантов последовательностей. Для генерации последовательности значений нормально распределенной случайной величины (2) нам необходимы две равномерно распределенные величины. Если для одной из них мы задаем стартовую точку в виде показания таймера, то стартовую точку для второй можно задавать как некоторую функцию от первой.

При задании различных таких зависимостей явления, сопровождающие построение нормально распределенной величины, могут меняться.

В качестве примера ограничимся рассмотрением случайной величины построенной по следующим параметрам, с использованием (2):

а) стартовая точка $X_0(\alpha_1)$ равна текущему значению секунд, снятому с системного таймера;

$$\text{б) } X_0(\alpha_2) = X_0(\alpha_1) + 1; \tag{6}$$

в) для α_1, α_2 $a = 9$, $c = 3$, $m = 1024$.

Для исследования построим выборки двух размеров – 1024 и 512 значений. Результаты исследований представлены в таблице.

Результаты исследования по χ^2

Метод	Количество экспериментов	Число реализаций случайной величины в каждом эксперименте	df _{мода}	χ^2 среднее	Р _{ср.}	σ_p
Метод (2)	60	1024	9	2,878	0,914	0,238
Метод (2)	60	512	7	5,25	0,646	0,269
STATISTICA	20	1024	9	7,944	0,5707	0,274
STATISTICA	20	512	7	8,011	0,446	0,311
RandG	40	1024	9	9,2526	0,4849	0,2885
RandG	40	512	7	7,399	0,457	0,292

Необходимо отметить, что при исследовании построенных по (2), (6) случайных величин на соответствие нормальному закону распределения по критерию χ^2 выявилась периодичность. На рис. 1 представлена зависимость p от стартовой точки $X_0(\alpha_1) = 1, 2, \dots, 61$ для 1024 реализаций построенной случайной величины, на рис. 2 – для 512 реализаций той же величины.

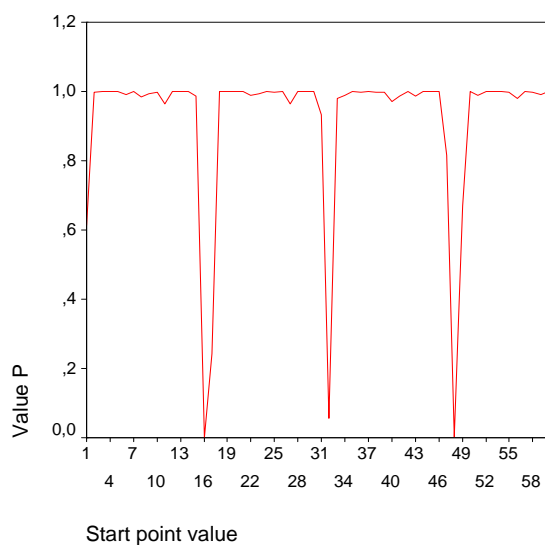


Рис.1

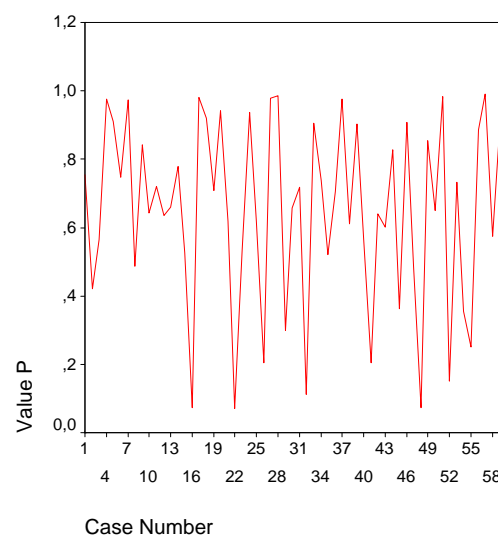


Рис. 2

Выводы

Рассмотрены различные методы моделирования нормально распределенных случайных величин. Подробно описана проблема использования метода (2) с целью его программной реализации. Исследованы различные методы на соответствие нормальному распределению при помощи критерия χ^2 .

При работе со случайными величинами трудно дать однозначные рекомендации по использованию того или иного метода. Требования к построенной величине могут варьироваться и зависят от специфики задачи. На основании проведенного исследова-

ния и его результатов, описанных в таблице, возможно выбрать наиболее удобный метод моделирования с требуемыми параметрами.

Литература

1. **Вуколов Э. А.** Основы статистического анализа. – М.: Форум–Инфра, 2004.
2. **Сушков Ю. А.** Статистические модели систем. – СПб.: изд-во СПбГУ, 2004.
3. Справочник по прикладной статистике/Под ред. Э. Ллойда, У.Ледермана. – М.: Финансы и статистика. Т. 1. – 1989; Т. 2. – 1990.
4. **Юров В. И.** Assembler практикум. – СПб.: Питер, 2004.
5. **Кнут Д.** Искусство программирования для ЭВМ. Т. 2. – М.: Мир, 1977.