11. Дослідження функції дихання [Електронний ресурс] / Ukrainian Context Optimizer. – Режим доступа: http://intranet.tdmu.edu.te.ua/data/cd/tuberkulez/html/Rozdil08/r08.html.

12. Спирометрия в оценке нарушений функции дыхательной системы [Електронний ресурс] / Ukrainian Context Optimizer. – Режим доступа: http://health-ua.com/articles/2426.html.

13. Спирометрия ХОБЛ [Електронний ресурс] / Ukrainian Context Optimizer. – Режим доступа: http://www.medicaljournalsworld.com/topics.

14. Оксиметрия [Електронний ресурс] / Ukrainian Context Optimizer. – Режим доступа: http://www.pediatr-site.ru/187-oksimetriya.html.

15. Калакутский Л.И. Аппаратура и методы клинического мониторинга / Л.И. Калакутский, Э.С. Маннелис– М. : Высш.шк., 2004. – 156 с.

16. Oximetrix 3 – Sv02 Sistem. Abbott Lab. Ltd. – CA, USA, 1990. – 6 p.

17. Церебральная оксиметрия (rSO2) [Електронний ресурс] / Ukrainian Context Optimizer. – Режим доступа: http://physiomed.com.ua/index.php/razdely-meditsiny/nevrologiya-i-nejrokhirurgiya/nejroreanimatsiya/132-nejromonitoring/524-cerebralnaya-oksimetriya-rso2.

18. Церебральный оксиметр INVOS-3100. – Somanetics Corp. Michigan, USA, 1990. – 6 p.

19. Функціональні методи дослідження в пульмонології [Електронний ресурс] / Ukrainian Context Optimizer. – Режим доступа: http://medstrana.com/articles/130.

20. Пневмотахометрия [Електронний ресурс] / Ukrainian Context Optimizer. – Режим доступа: http://www.spontan.ru/spravochnik-pulmonologa/504-pnevmotaxometriya.html.

21. Клиническая капнография [Електронний ресурс] / Ukrainian Context Optimizer. – Режим доступа: http://www.basko.spb.ru/article_37.html.

22. Тінтіналлі Дж. Е. Невідкладна медична допомога / Тінтіналлі Дж. Е., Кроума Рл., Руїза Е. – М.: Медицина, 2001. – 426 с.

23. Малишев В.Д. Інтенсивна терапія. Реанімація. Перша допомога: [навчальний посібник] / Малишев В.Д. – М. : Медицина, – 2000. – 464 с.

24. Капнометрия и капнография (капнография в картинках) [Електронний ресурс] / Ukrainian Context Optimizer. – Режим доступа: http://rusanesth.com/spezialistam/rukovodstva/3.html.

25. Medley W. Noninvasive Blood Gas Monitoring // Clinical Blood Gases. 1990. 281 – 301 p.

Надійшла 18.9.2012 р.
Рецензент: д.т.н. Петрук В.Г.

УДК 621.391:004.73

Э.В. ФАУРЕ, А.С. БЕРЕЗА, Е.А. ЯРОСЛАВСЬКАЯ
Черкасский государственный технологический университет

ОЦЕНКА ТОЧНОСТИ ВОСПРОИЗВЕДЕНИЯ ЗАКОНА РАСПРЕДЕЛЕНИЯ ДИСКРЕТНОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ ПРИ ЕЁ ПРЕОБРАЗОВАНИИ

Рассматриваются вопросы количественной оценки ошибки воспроизведения закона распределения дискретной случайной величины при изменении ее области определения. Показано, что если первичный генератор порождает равномерно распределенную дискретную случайную величину с нулевой ошибкой воспроизведения, то после преобразования случайной величины возникает ошибка воспроизведения ее закона распределения. Решена задача минимизации ошибки воспроизведения закона распределения дискретной случайной величины при изменении ее области определения.

The problems of estimating the reproduction error value of the distribution of a discrete random variable with a change in its scale are reviewed in the article. It is shown that if the primary generator generates uniformly distributed discrete random variables with zero-error reproduction law of distribution, then reproduction error of the distribution of the random variable appears after the transformation of the scale of the random variables at the output of the converter. Questions about minimization of reproduction error of the distribution of a discrete random variable after changing its scale are solved.

Ключевые слова: генератор случайных чисел, область определения случайной величины, ошибка воспроизведения закона распределения.

Введение

Многие практические задачи, в частности, задачи, связанные с криптографией или имитационным моделированием [1– 3], не могут быть решены без применения равномерно распределенных случайных последовательностей. В силу этого обстоятельства большинство компьютерных программ включают в себя процедуру генерации случайных чисел, за которой закрепилось название random. Для решения этой задачи программа random создает генератор равномерно распределенных на отрезке $[0, N_1 - 1]$ чисел, при этом величина N_1 определяется разрядностью процессора. Для основной массы ныне используемых коммерческих ПЭВМ $N_1 = 2^n$ где $n = 32$. При практическом использовании генератора случайных чисел (ГСЧ) область определения равномерно распределенной случайной величины определяется конкретной

задачей и может отличаться от множества целых чисел отрезка $[0, N_1 - 1]$ – например, включать все целые числа отрезка $[0, N_2 - 1]$, где $N_2 < N_1$. В этом случае программа random сначала преобразует область определения случайной величины из множества целых чисел отрезка $[0, N_1 - 1]$ к множеству рациональных чисел полуинтервала [0, 1) путем выполнения операции $\hat{x} = \frac{x}{N_1}$, а затем вычисляет $\hat{y} = \hat{x} N_2 = \left\lfloor x \frac{N_2}{N_1} \right\rfloor$, при этом $0 < \hat{x} < 1$. Здесь выражение $\lfloor A \rfloor$ означает целую часть (функцию пол) числа А. Такая практика вычислений используется повсеместно, хотя вопрос точности воспроизведения случайных величин остается неизученным и, как следствие, вопрос точности конечного результата, обусловленный наличием ошибки воспроизведения закона распределения, остается открытым.

Выделение не решенных ранее частей общей проблемы

К настоящему времени получен большой опыт генерации равномерно распределенных чисел и их статистической обработки, однако вопрос точности воспроизведения закона распределения остается недостаточно изученным, несмотря на то, что теория математической статистики дает ясные ответы на многие вопросы, связанные с интервальными и точечными статистиками, оценками статистических гипотез и т.п. [4, 5]. В частности, для оценки точности воспроизведения законов распределения дискретной случайной величины математическая статистика использует функцию ошибки, определяемую как

$$\xi(x) = p_0(x) - \hat{p}(x). \tag{1}$$

Вероятность

$$p_0(x) = \frac{1}{N_1} \tag{2}$$

соответствует гипотетическому (теоретическому) закону распределения равномерно распределенной случайной величины *x* на области ее определения $(x \in [0, N_1 - 1])$,

$$\hat{p}(x) = \frac{n(x)}{V}. \tag{3}$$

соответствует эмпирическому (экспериментально полученному) закону распределения случайной величины в той же области, *n* (x) – число случаев появления числа x в выборке объема *V*.

Функция (1) лежит в основе критерия Колмогорова-Смирнова (КС-критерия) и критерия Пирсона (χ^2-критерий). По КС-критерию оценивается максимальное значение выражения (1):

$$\xi_{\max}(x) = \max \left| p_0(x) - \hat{p}(x) \right|, \tag{4}$$

по которому делается вывод, является ли исследуемая выборка выборкой из генеральной совокупности равномерно распределенной дискретной случайной величины или нет.

По критерию Пирсона оценивается величина

$$\chi^2 = V \sum_{x=0}^{N_1 - 1} \frac{(p_0(x) - \hat{p}(x))^2}{p_0(x)}, \tag{5}$$

Эта величина сравнивается с квантилью закона распределения χ^2 заданного уровня значимости α. Критерий удобен для выборок сравнительно малых объемов и сравнительно больших значений ошибки воспроизведения. В настоящей работе рассмотрим случай больших объемов выборок, больших значений *N* и нулевой (или очень близкой к нулю) ошибки воспроизведения закона распределения дискретной случайной величины.

Постановка задачи

Задачей данного исследования является определение величины ошибки воспроизведения дискретной случайной величины при изменении области ее определения от множества целых чисел отрезка $[0, N_1 - 1]$ к множеству целых чисел отрезка $[0, N_2 - 1]$, $N_2 < N_1$.

Решение задачи

В данной работе вопрос проверки гипотез обсуждаться не будет – полагается, что выборка объема *V* принадлежит генеральной совокупности равномерно распределенной дискретной случайной величины. Определяется величина ошибки воспроизведения закона распределения дискретной случайной величины.

Для установления степени соответствия эмпирического закона теоретическому рассмотрим более детально выражение (1), для чего представим его в виде

$$\xi(x) = p_0(x) - \hat{p}(x) = \frac{n_0(x)}{V} - \frac{n(x)}{V} = \frac{\Delta n(x)}{V}, \tag{6}$$

где $\Delta n(x) = n_0(x) - n(x)$,

$n_0(x)$ – количество повторений символа *x* в теоретическом потоке;

$n(x)$ – количество повторений символа *x* в эмпирическом потоке.

Полученное выражение допускает следующую трактовку протекающего процесса: дискретный случайный процесс, порождаемый реальным ГСЧ, представляет композицию двух дискретных случайных процессов. Первый из них – это поток символов с теоретическим законом распределения, а второй – это поток мешающих символов с неизвестным законом распределения. Соответственно, каждый из потоков порождается своим генератором. Генератор мешающих символов может вставлять порождаемые им символы в общий поток и может подавлять (удалять) символы теоретического потока. Вставка/подавление символов трансформирует теоретическое распределение в эмпирическое. Тогда ошибка воспроизведения закона распределения случайной величины (6) есть число символов мешающего потока, поразивших некоторый символ, приходящееся на единицу объема выборки. С учетом этого сформулируем критерий оценки точности воспроизведения закона распределения дискретной случайной величины реальным ГСЧ:

$$\xi = \frac{1}{2}\sum_{x=0}^{N_1-1}\xi(x) = \frac{1}{2}\sum_{x=0}^{N_1-1}\frac{|\Delta n(x)|}{V} = \frac{1}{2V}\sum_{x=0}^{N_1-1}|\Delta n(x)|, \qquad (7)$$

которым и будем пользоваться в данной работе. Значение 2 дополнительно вводится в знаменатель в силу того, что один символ мешающего потока изменяет статистику повторений сразу для двух символов алфавита.

Выражение (7) означает, что точность воспроизведения закона распределения дискретной случайной величины реальным генератором определяется числом символов ошибочного потока, приходящимся на единицу объема выборки.

Обратим внимание на то, что если объем выборки не кратен N_1, то появляется дополнительно статистическая составляющая ошибки воспроизведения закона распределения дискретной случайной величины. Это обусловлено тем, что значение $n_0 = \frac{V}{N_1}$ является целым в том и только в том случае, если $V = kN_1$, а статистическая ошибка – ошибка, определяемая степенью отклонения объема выборки от величины $V = kN_1$, – равна нулю.

Неравенство нулю выражения (7) на промежутке, кратном периоду формируемой последовательности, характеризует конструктивную ошибку воспроизведения закона распределения дискретной случайной величины случайного процесса, порожденного реальным ГСЧ (или преобразователем) – ошибку, обусловленную принципом построения (конструкцией) ГСЧ или преобразователя.

Определим величину статистической ошибки (динамической составляющей потока мешающих символов), обусловленной некратностью объема выборки циклу ГСЧ. Под циклом ГСЧ будем понимать последовательность на его выходе, которая циклически повторяется при неизменных параметрах ГСЧ.

Пусть последовательность на выходе ГСЧ состоит из циклов длины N_1, которые в случайном порядке содержат все значения случайной величины из области определения, а порядок следования символов в циклах может отличаться. Объем выборки $V = kN_1 + m$, где $0 < m < N_1$ – число символов последнего (неполного) цикла, вошедшего в выборку.

Тогда $n_0 = \frac{kN_1 + m}{N_1} = k + \varepsilon$, где $\varepsilon = \frac{m}{N_1}$. Учтем, что при любом V $p_0(x) = \frac{n_0}{V} = \frac{V/N_1}{V} = \frac{1}{N_1}$, поэтому

$\xi(x) = p_0(x) - \hat{p}(x) = \frac{n_0(x)}{V} - \frac{n(x)}{V} = \frac{1}{N_1} - \frac{n(x)}{V}$. Значение статистической ошибки воспроизведения $\xi = \frac{1}{2}\sum_{x=0}^{N_1-1}\xi(x) = \frac{1}{2}\left((N_1 - m)\left|\frac{1}{N_1} - \frac{k}{V}\right| + m\left|\frac{1}{N_1} - \frac{k+1}{V}\right|\right) = \frac{m(N_1 - m)}{N_1 V}$ или

$$\xi = \frac{\varepsilon(1-\varepsilon)}{k+\varepsilon}. \qquad (8)$$

Построим график зависимости величины ошибки воспроизведения от объема выборки.

Из графика рис. 1 видно, что при условии равномерного распределения слов из всей области определения случайной величины в каждом из циклов статистическая ошибка воспроизведения закона распределения дискретной случайной величины имеет нулевое значение при $\varepsilon = 0$ и $\varepsilon = 1$ (в точках $V = kN_1$) и уменьшается с увеличением объема выборки.

В качестве условия пригодности ГСЧ (или метода преобразования случайной последовательности чисел) будем использовать условие $\xi \le 10^{-3}$, что означает: ГСЧ или преобразователь пригоден для

практического применения, если на каждую тысячу символов потока приходится не более одного мешающего символа.

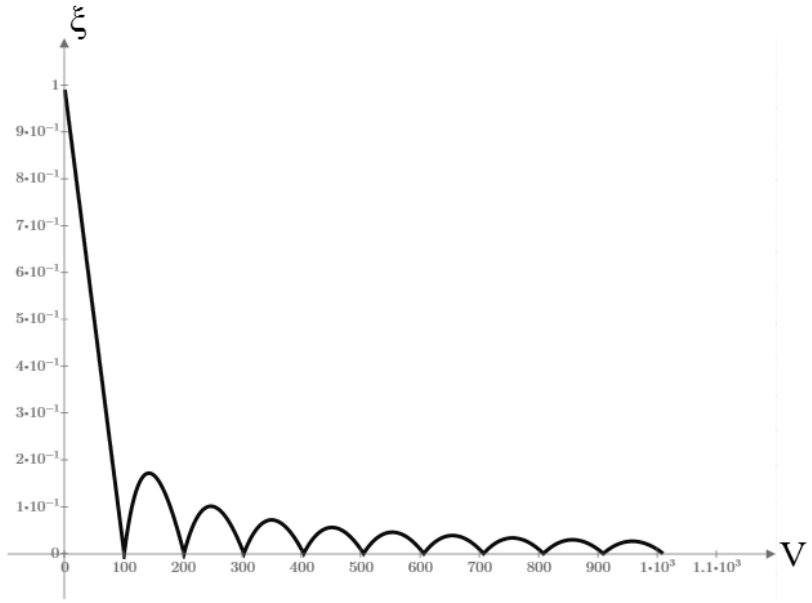

Рис. 1. График зависимости величины ошибки воспроизведения от объема выборки при N_1=101

С учетом предложенного критерия (7) рассмотрим свойства последовательностей и ошибку воспроизведения для двух наиболее распространенных типов ГСЧ:

- генератора М-последовательности на регистрах сдвига с обратными связями;
- генератора конгруэнтной последовательностей чисел (ГКЧ).

Известно, что генераторы М-последовательностей порождают периодически повторяемую псевдослучайную последовательность чисел длиной $N_1 = 2^n - 1$, где n – порядок генераторного полинома. Порождаемая этим генератором последовательность содержит все числа отрезка $[0, N_1 - 1]$. Отсюда следует, что для области определения случайной величины $x \in [1, N_1 - 1]$ генератор М-последовательности порождает равномерно распределенную случайную последовательность чисел. Для $x \in [0, N_1 - 1]$ этот же генератор порождает случайную величину с ошибкой, обусловленной отсутствием в последовательности символа «0». Теоретическое и эмпирическое распределение такого генератора на объеме выборки $V = N_1$, а также его ошибка воспроизведения показаны на рис. 2.

Для этого генератора в соответствии с предложенным критерием (7) получим, что при $V = N_1$ ошибка, обусловленная конструкцией генератора, $\xi = \dfrac{\sum_{x=0}^{N-1} |\Delta n(x)|}{2V} = \dfrac{1}{N_1}$.

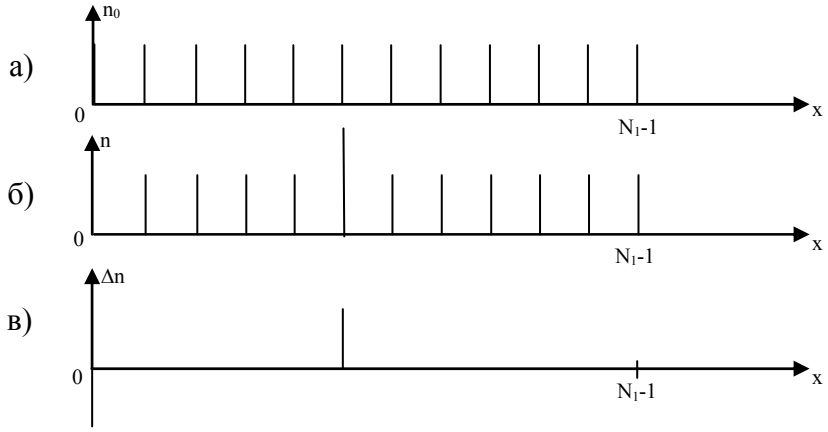

Рис. 2. Законы распределения составляющих полного потока символов на выходе генератора М-последовательности: а) теоретического потока; б) эмпирического потока; в) ошибки воспроизведения

Из этого следует, что при $N_1 > 1000$ (порядок генераторного полинома $n \geq 10$) генератор М-

последовательности удовлетворяет приведенному требованию $\xi \le 10^{-3}$. Граф состояний этого генератора имеет вид $\Gamma = 1 \times (N_1 - 1) + 1 \times 1$, т. е. содержит один цикл длиной $N_1 - 1$ и один нуль-цикл (цикл длиной 1), расположенный в точке $x = 0$.

Известен способ рандомизации случайных последовательностей, порождаемых генератором конгруэнтных чисел [6], сущность которого заключается в конкатенации всех циклов графа состояний. Практически это означает, что если цикл генератора М-последовательности дополнить словом $x = 0$ в произвольно выбранном месте последовательности, то период последовательности на выходе генератора станет равной N_1, а ошибка воспроизведения $\xi = 0$.

Для ГКЧ величина ошибки определяется конструкцией графа состояний. Так, например, для конструкции графа $\Gamma = 1 \times N_1$ величина ошибки воспроизведения будет равна нулю, Однако в этом случае символы последовательности имеют наибольшую степень взаимной корреляции, т.к. порядок чередования слов в последовательности является детерминированным. При конструкции графа ГКЧ $\Gamma = 1 \times (N_1 - 1) + 1 \times 1$ (совпадающей с конструкцией графа состояний генератора М-последовательности) величина ошибки и степень корреляции между символами ГКЧ совпадают с величиной ошибки и степенью корреляции между символами М-последовательности. Отличие заключается лишь в том, что у генератора М-последовательности отсутствующим символом всегда является символ «0», а у ГКЧ нуль-цикл может располагаться в произвольной точке отрезка $[0, N_1 - 1]$, значение этого символа определяется параметрами ГКЧ, т.е. его конструкцией.

Для ГКЧ с графом состояний типа $\Gamma = d \times t + 1 \times 1$, где *d* – число циклов длиной *t* символов каждый, величина ошибки воспроизведения может быть очень большой и определяется длиной цикла. Так, для $N_1 = 100$ и конструкции графа $\Gamma = 5 \times 20 + 1 \times 1$ число различных символов в последовательности длины $V = N_1$, состоящей из символов некоторого цикла, отличного от нуль-цикла, будет равняться 20, а $\xi = \dfrac{\sum_{x=0}^{N-1} |\Delta n(x)|}{2V} = \dfrac{19 \cdot 4 + 1 \cdot 5 + 81 \cdot 1}{2 \cdot 101} = \dfrac{81}{101} \approx 0{,}8$. Такой генератор едва ли можно назвать генератором равномерно распределенной на отрезке $[0, N_1 - 1]$ случайной последовательности чисел. Однако ситуация в корне меняется, если выполнить конкатенацию циклов графа. В этом случае ошибка воспроизведения $\xi = 0$, а корреляция между символами будет существенно ниже. Далее будем исходить именно из этой ситуации – первичный ГСЧ порождает последовательность случайных чисел отрезка $[0, N_1 - 1]$, а ошибка воспроизведения равномерного закона распределения случайной величины $\xi = 0$.

Положим, что необходимо сформировать равномерно распределенную случайную последовательность чисел отрезка $[0, N_2 - 1]$, где $N_1 < N_2$, причем все значения из области определения случайной величины в сформированной последовательности должны встречаться один раз. Положим также, что

$$1 \le \frac{N_1 - 1}{N_2 - 1} < 2 . \qquad (9)$$

Отметим, что если $\dfrac{N_1 - 1}{N_2 - 1} \ge 2$, то выполняется процедура прореживания столько раз, сколько нужно для получения соотношения (9). Прореживание включает следующие операции:

- из ряда генерируемых чисел удаляются нечетные числа (остаются только четные – это, собственно, и есть прореживание);
- разделив оставшиеся четные числа на два, формируется натуральная последовательность чисел, равномерно распределенных на отрезке $\left[0, \dfrac{N_1 - 1}{2}\right]$ при N_1 – нечетном и $\left[0, \dfrac{N_1 - 2}{2}\right]$ при N_1 – четном.

Заметим, что при $N_1 - 1 = 2^k (N_2 - 1) + l$, $l \in \left\lfloor 0, 2^k - 1 \right\rfloor$, и выполнении операции прореживания распределение случайной величины на отрезке $[0, N_2 - 1]$ будет равномерным.

Рассмотрим два способа изменения области определения случайной величины:

1. $y = \left\lfloor x \dfrac{N_2}{N_1} \right\rfloor$, выражение $\lfloor A \rfloor$ означает целую часть (функцию пол) числа А;
2. $y = |x|_{N_2}$, выражение $|A|_B$ означает вычет числа *A* по модулю *B*.

Найдем величину ошибки воспроизведения закона распределения случайной величины *y* для первого и второго способа ее формирования, для чего рассмотрим фрагмент исходной последовательности

(в качестве которой используем натуральную последовательность чисел 0,1,2,3…., начиная с числа 143) и выходную последовательность для каждого из способов преобразования. Результаты сведем в таблицу. $N_1 = 256$, $N_2 = 151$.

Таблица 1

Фрагменты исходной и преобразованной последовательностей

x	143	144	145	146	147	148	149	150	151	152	153
$y = \left\lfloor x\frac{N_2}{N_1} \right\rfloor = \left\lfloor x\frac{151}{256} \right\rfloor$	84	84	85	86	86	87	87	88	89	89	90
$y = \lvert x \rvert_{N_2} = \lvert x \rvert_{151}$	143	144	145	146	147	148	149	150	0	1	2

Из таблицы видно, что для первого способа преобразования некоторые слова вследствие возникновения ситуаций $\left\lfloor x\frac{N_2}{N_1} \right\rfloor = \left\lfloor (x+1)\frac{N_2}{N_1} \right\rfloor$ появляются на выходе преобразователя дважды. Число слов-двойников на выходе преобразователя после поступления на его вход N_1 символов первичного генератора равняется $\Delta n = N_1 - N_2$. Ошибка воспроизведения закона распределения случайной величины на выходе преобразователя $\xi = \frac{1}{2}\sum_{x=0}^{N_2-1}\xi(x) = \frac{1}{2}\left((N_1 - N_2)\left|\frac{1}{N_2} - \frac{2}{N_1}\right| + (2N_2 - N_1)\left|\frac{1}{N_2} - \frac{1}{N_1}\right|\right)$ или

$$\xi = \frac{(N_1 - N_2)(2N_2 - N_1)}{N_1 N_2}. \qquad (10)$$

Для рассмотренного примера $\xi = \frac{(N_1 - N_2)(2N_2 - N_1)}{N_1 N_2} = \frac{(256-151)(2\cdot 151 - 256)}{256\cdot 151} \approx 0{,}12$, что ставит под сомнение целесообразность этого преобразования. Из полученного результата также следует, что число ошибочных символов на интервале из $N_1 N_2$ слов равняется $(N_1 - N_2)(2N_2 - N_1)$.

Следует отметить, что поскольку первичный ГСЧ циклически повторяет последовательность чисел отрезка $[0, N_1 - 1]$, то и преобразователь $y = \left\lfloor x\frac{N_2}{N_1} \right\rfloor$ будет циклически повторять выходную последовательность с одинаковым количеством слов-двойников в каждом цикле. Отсюда следует, что ошибка

$$\xi = \frac{1}{2}\sum_{x=0}^{N_2-1}\xi(x) = \frac{1}{2}\left((N_1 - N_2)\left|\frac{1}{N_2} - \frac{2k}{kN_1}\right| + (2N_2 - N_1)\left|\frac{1}{N_2} - \frac{k}{kN_1}\right|\right) = \frac{(N_1 - N_2)(2N_2 - N_1)}{N_1 N_2}$$

сохраняется при любых объемах выборки $V = kN_1$ и определяет конструктивную ошибку преобразователя $y = \left\lfloor x\frac{N_2}{N_1} \right\rfloor$. При $V = kN_1 + m$, помимо конструктивной составляющей погрешности, возникает статистическая составляющая погрешности.

Для второго способа преобразования $y = \lvert x \rvert_{N_2}$. Таким образом, при $x < N_2$ $y = x$, а при $x \ge N_2$ $y = x - N_2$.

Поскольку числа x и $x + N_2$ имеют равные остатки при делении на N_2, количество слов, равное $\Delta n = N_1 - N_2$, в последовательности будет встречаться дважды. Вследствие этого, ошибка воспроизведения такого преобразователя будет также определяться выражением (10).

Отсюда следует вывод, что ни один из рассмотренных способов преобразования не удовлетворяет поставленным требованиям и, в общем, непригоден для практического использования.

Таким образом, способ изменения области определения случайной величины, обеспечивающий нулевую ошибку воспроизведения равномерного закона распределения, может включать следующие операции:

- выполнение процедуры прореживания исходной последовательности, чтобы выполнялось условие $1 \le \frac{N_1 - 1}{N_2 - 1} < 2$;
- выполнение преобразования $y = f(x)$ в соответствии с любым из рассмотренных способов;

- при необходимости прореживание полученного потока (удаление слов-двойников).

Поясним смысл операции прореживания. Поскольку первичный генератор чисел *x* порождает слова отрезка $[0, N_1 - 1]$, а вторичный генератор слов *y* должен формировать слова отрезка $[0, N_2 - 1]$, то $\Delta n = N_1 - N_2$ слов из N_1 слов на выходе вторичного ГСЧ являются избыточными и подлежат удалению. Эта операция достаточно просто выполняется при $y = |x|_{N_2}$. В этом случае подлежат удалению с входа преобразователя все слова $x \geq N_2$, которые и порождают слова-двойники.

Для $y = \left\lfloor x \frac{N_2}{N_1} \right\rfloor$ подлежат удалению по одному из пары слов, удовлетворяющих условию $y(x+1) - y(x) = 0$. Эта операция легко выполнима, если слова $y(x)$ и $y(x+1)$ следуют одно за другим. В общем случае, если первичный ГСЧ выдает некоррелированную последовательность чисел, это условие не выполняется, что приводит к практической неэффективности этого метода для решения поставленной задачи.

Полученные результаты

Проведенное исследование показывает, что преобразование области определения случайной величины из множества целых чисел отрезка $[0, N_1 - 1]$ в множество целых чисел отрезка $[0, N_2 - 1]$ при объеме исходной выборки, кратном N_1, где $1 \leq \frac{N_1 - 1}{N_2 - 1} < 2$, путем вычисления функции $y = \left\lfloor x \frac{N_2}{N_1} \right\rfloor$ или путем вычисления функции $y = |x|_{N_2}$ приводит к появлению конструктивной ошибки преобразования, равной $\xi = \frac{(N_1 - N_2)(2N_2 - N_1)}{N_1 N_2}$. При $N_1 - 1 = 2^k (N_2 - 1) + l$, $l \in [0, 2^k - 1]$, и выполнении операции прореживания распределение случайной величины на отрезке $[0, N_2 - 1]$ будет равномерным с нулевой конструктивной ошибкой преобразования.

Величина статистической ошибки (динамической составляющей потока мешающих символов), обусловленной некратностью объема выборки циклу ГСЧ длины N_1, определяется выражением $\xi = \frac{1}{2} \sum_{x=0}^{N_1 - 1} \xi(x) = \frac{m(N_1 - m)}{N_1 V} = \frac{\varepsilon(1 - \varepsilon)}{k + \varepsilon}$, где $V = kN_1 + m$, $0 < m < N_1$, а $\varepsilon = \frac{m}{N_1}$.

Выводы

Преобразование области определения равномерно распределенной дискретной случайной величины с нулевой ошибкой закона распределения может быть обеспечено совместным применением преобразования $y = |x|_{N_2}$ с последующим удалением слов-двойников.

Литература

1. Хеерман Д. В. Методы компьютерного эксперимента в теоретической физике / Хеерман Д. В. – М. : Наука, 1990. – 176 с.

2. Соболь И. М. Метод Монте-Карло / Соболь И. М. – М. : Наука, 1985. – 80 с.

3. Диффи У. Защищенность и имитостойкость: Введение в криптографию / У. Диффи, М. Э. Хеллман // ТИИЭР, 1979. – Т. 67. – № 3. – С. 71– 109.

4. Боровков А. А. Математическая статистика / Боровков А. А. – М. : Наука, 1984. – 472 с.

5. Ивченко Г. И. Математическая статистика / Г. И. Ивченко, Ю. И. Медведев – М. : Высшая школа, 1984. – 284 с.

6. Пат. 41079 Україна, МПК G06F 7/58. Спосіб рандомізації послідовності конгруентних чисел / Мітянкіна Т.В.; Швидкий В.В.; Щерба А.І.; Мітянкін М.О.; заявник та патентовласник Черкаський державний технологічний університет – № u200808187; заявл. 17.06.2008; опубл. 12.05.2009, Бюл. № 9.

Надійшла 12.9.2012 р.
Рецензент: д.т.н. Рудницький В.М.