

УДК 519.9

В.Ю. ДУБНИЦКИЙ¹, А.Г. ПРОЦЕНКО²

¹Харьковский институт банковского дела Университета банковского дела, Национального банка Украины, Украина

²Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Украина

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ДАТЧИКОВ СЛУЧАЙНЫХ ЧИСЕЛ СИСТЕМ STATGRAPHICS И MATHCAD

Выполнено исследование датчиков случайных чисел систем STATGRAPHICS и MATHCAD. Показано, что при любом количестве данных совпадение законов распределения и их параметров отмечено только для равномерного распределения. Для других законов распределения случайных величин совпадение отмечено не во всех случаях даже при генерировании выборки объемом в тысячу данных. Желательно проведение дополнительных исследований задачи оценки качества случайных чисел, полученных различными системами моделирования.

Ключевые слова: STATGRAPHICS, MATHCAD, датчики случайных чисел, качество датчиков случайных чисел, законы распределения.

Введение

При имитационном моделировании технических, организационных и иных систем широко используют датчики случайных чисел (ДСЧ). Результаты могут быть использованы как самостоятельно, так и в качестве исходных данных для дальнейших исследований. Правила построения ДСЧ изложены в классической работе [1].

Для наиболее распространенных законов распределения способы генерирования случайных чисел приведены в работе [2]. В работе [1] приведено описание алгоритма, проверяющего качество ДСЧ для получения данных, подчиняющихся равномерному закону распределения.

Качество ДСЧ будем оценивать по двум характеристикам:

1. Совпадение предполагаемого и фактически полученного законов распределений;
2. Совпадение предполагаемых и фактически полученных параметров этих законов.

Таким образом, возникает задача оценки качества ДСЧ входящих в состав данной системы и сравнительная оценка ДСЧ, входящих в различные системы. Появление этой задачи обусловлено тем, что в случае обмена данными между различными программными продуктами требуется уверенность в их однозначной интерпретации.

Авторам не удалось найти в доступной им информации сведения об аналогичных исследованиях.

Целью данной работы было сравнение ДСЧ для различных законов распределения, входящих в системы STATGRAPHICS и MATHCAD.

1. Методика получения экспериментальных данных

В качестве объектов исследования были выбраны системы MATHCAD 2001 [3] и STATGRAPHICS V.15 [4].

Собственно исследование состояло из нескольких этапов:

1. В системе MATHCAD генерировали данные, соответствующие следующим законам распределения: равномерному (ЗН1), нормальному (ЗН2), логнормальному (ЗН3), гамма-распределению (ЗН4), бета-распределению (ЗН5), распределению Коши (ЗН6), распределению Лапласа (ЗН7), экспоненциальному (ЗН8), логистическому (ЗН9), распределению Вейбула (ЗН10).

2. Для каждого из этих законов генерировались выборки объемом 10, 50, 100, 500, 1000 данных.

3. Полученные одномерные выборки случайных чисел проверяли на соответствие выбранным законам распределения. Проверка этих законов проводилась с помощью средств пакета STATGRAPHICS.

4. В системе STATGRAPHICS генерировались данные, соответствующие законам распределения: указанным в п.1. Для каждого из этих законов также генерировали выборки объемом 10, 50, 100, 500, 1000 данных.

5. Полученные одномерные выборки случайных чисел проверялись на соответствие выбранным законам распределения с помощью средств той же системы STATGRAPHICS.

Целью последних двух этапов была проверка способности системы STATGRAPHICS правильно

распознать ею же сгенерированные случайные числа. Структурно-логическая схема исследования показана на рис. 1.



Рис. 1. Структурно-логическая схема исследования качества ДСЧ

2. Изложение результатов

Перечень выбранных законов распределения целиком обусловлен возможностями системы MATHCAD. Все эти законы широко используют

при моделировании надёжности систем различной природы.

Свойства этих законов описаны в работе[2]. Результаты проверки качества датчиков случайных чисел приведены в табл.1. Из этой таблицы следует, что из пятидесяти вариантов данных, сгенерированных системой MATHCAD, только в четырнадцати случаях (28%) система STATGRAPHICS верно распознала закон распределения. Анализ данных, приведенных в табл.1 позволил сделать следующие выводы.

Система STATGRAPHICS верно распознаёт данные, сгенерированные системой MATHCAD независимо от объёма, если они равномерно распределены. Для моделирования наиболее распространённых в практике законов распределения: нормально-го, логарифмически нормального, Вейбулла минимальный удовлетворительный объём данных начинается от ста случайных чисел.

Следует отметить, что такой важный для практики вид закона, как экспоненциальный не был правильно распознан при всех исследованных объёмах выборок.

Таблица 1

Результат проверки данных, сгенерированных пакетом MATLAB на соответствие требуемому закону распределения средствами пакета STATGRAPHICS

| Условные обозначения данных, сгенерированных в пакете MATHCAD в соответствии с требуемым законом распределения*) | Результат проверки данных средствами пакета STATGRAPHICS | | | | |
|--|--|---|---|-------|-------------------------------|
| | Объём выборки | | | | |
| | 10 | 50 | 100 | 500 | 1000 |
| ЗН1 | + | + | + | + | + |
| ЗН2 | ЗН1 | ЗН9 | + | + | + |
| ЗН3 | ЗН1 | Распределение наибольшего экстремального значения | Распределение наибольшего экстремального значения | ЗН4 | Обратное распределение Гаусса |
| ЗН4 | ЗН1 | Распределение Бирнбаума-Сандерса | + | + | + |
| ЗН5 | отказ**) | отказ | отказ | отказ | отказ |
| ЗН6 | отказ | отказ | отказ | отказ | отказ |
| ЗН7 | ЗН1 | ЗН1 | ЗН10 | ЗН1 | ЗН1 |
| ЗН8 | ЗН1 | ЗН10 | ЗН4 | ЗН10 | ЗН4 |
| ЗН9 | отказ | отказ | отказ | отказ | отказ |
| ЗН10 | ЗН1 | ЗН1 | + | + | + |

*) – названия законов соответствуют указанным в разделе «Методика получения экспериментальных данных», пункт 1.

***) – отказ системы STATGRAPHICS от распознавания закона распределения.

При распознавании вида закона распределения наилучшим образом описывающего сгенерированные случайные числа пакет STATGRAPHICS выделил наряду с уже названными также сравнительно редко встречающиеся распределения: распределение Бирнбаума-Сандерса и обратное распределение Гаусса. Рассмотрим их свойства более подробно. Свойства распределения Бирнбаума-Сандерса рас-

смотрены в работе [5]. Это распределение было введено в 1969г. для описания результатов усталостных испытаний.

Плотность распределения Бирнбаума-Сандерса имеет вид:

$$f(t, \theta, \beta) = \frac{\sqrt{\frac{t}{\theta}} + \sqrt{\frac{\theta}{t}}}{2\beta t} \phi(z), t > 0, \theta > 0, \beta > 0, \quad (1)$$

где $\phi(z)$ – плотность нормального закона распределения. Аргумент

$$z = \frac{1}{\beta} \left(\sqrt{\frac{t}{\theta}} - \sqrt{\frac{\theta}{t}} \right). \quad (2)$$

Математическое ожидание M и дисперсия σ^2 как функции аргументов-параметров закона распределения θ и β выражены так:

$$M = \theta \left(1 + \frac{\beta^2}{2} \right); \quad (3)$$

$$\sigma^2 = (\theta\beta)^2 \left(1 + \frac{5\beta^2}{4} \right). \quad (4)$$

Необходимые сведения об обратном распределении Гаусса приведены в работе [6].

Проверка способности системы STATGRAPHICS правильно распознать ею же сгенерированные

случайные числа, соответствующие указанным выше законам распределения, проводили так, как это показано на правой ветви рис.1.

Вначале генерировали одну тысячу случайных данных, соответствующих указанным выше законам распределения. Затем, с помощью системы EXCEL осуществляли случайную выборку необходимого объёма. На полученной выборке проводили распознавание закона распределения средствами пакета STATGRAPHICS. Результаты этого показаны в табл. 2. Все условные обозначения в табл.2 аналогичны табл. 1

Справочные сведения о логлогистическом распределении приведены в работе [5].

Для статистического анализа качества распознавания законов распределения сведём данные результатов распознавания из табл. 1 и 2 в табл.3.

Таблица 2

Результат проверки данных, сгенерированных пакетом STATGRAPHICS на соответствие требуемому закону распределения средствами пакета STATGRAPHICS

| Условные обозначения данных, сгенерированных в пакете STATGRAPHICS в соответствии с требуемым законом распределения | Результат проверки данных средствами пакета STATGRAPHICS | | | | |
|---|--|--------------------------------|-------------------------------|-------|----------------------------------|
| | Объём выборки | | | | |
| | 10 | 50 | 100 | 500 | 1000 |
| ЗН1 | + | + | + | + | + |
| ЗН2 | отказ | отказ | отказ | отказ | отказ |
| ЗН3 | ЗН1 | Обратное распределение Гаусса | Обратное распределение Гаусса | ЗН4 | Распределение Бирнбаума-Сандерса |
| ЗН4 | ЗН1 | Логлогистическое распределение | ЗН10 | + | + |
| ЗН5 | ЗН1 | ЗН1 | + | + | + |
| ЗН6 | Обратное распределение Гаусса | отказ | отказ | отказ | отказ |
| ЗН7 | ЗН1 | отказ | отказ | отказ | отказ |
| ЗН8 | Распределение Бирнбаума-Сандерса | ЗН4 | ЗН10 | ЗН10 | ЗН10 |
| ЗН9 | ЗН1 | + | ЗН7 | отказ | отказ |
| ЗН10 | ЗН1 | + | ЗН4 | ЗН4 | ЗН4 |

Таблица 3

Статистический анализ качества распознавания законов распределения

| Оценка распознавания законов распределения | Генерация данных в MATHCAD и проверка средствами STATGRAPH | Генерация данных в STATGRAPHICS и проверка собственными средствами пакета |
|--|--|---|
| Распознано верно | 14 | 12 |
| Распознано неверно | 15 | 23 |
| Отказ от распознавания | 21 | 15 |

Приведённые в табл. 3 данные исследованы в соответствии с методикой анализа таблиц Кх2. Методика этого анализа подробно описана в работе [7].

Так как для приведенных данных

$$\left(\chi_{\text{факт}}^2(2;0,05) = 2,84 \right) < \left(\chi_{\text{табл}}^2(2;0,05) = 5,99 \right),$$

то различие в качестве распознавания системой STATGRAPHICS данных, генерируемых системой MATHCAD, и распознавание системой STATGRAPHICS своих же данных отсутствует.

Выводы

1. Установлено, что специализированная статистическая система STATGRAPHICS в большинстве случаев не распознаёт соответствия заданному закону распределения случайных чисел, генерируемому системой MATHCAD.

2. Установлено, что специализированная статистическая система STATGRAPHICS в большинстве случаев не распознаёт соответствия заданному закону распределения случайных чисел, генерируемому ею же самой.

3. При взаимодействии различных систем компьютерной математики нежелательно проводить обмен данными, полученными в результате статистического моделирования, без специального исследования.

4. Желательно проведение дополнительных исследований задачи оценки качества случайных чисел, полученных различными системами моделирования.

Литература

1. Кнут Д. Искусство программирования в 3т. Т.2 : Получисленные методы / Д. Кнут. – М.: Вильямс, 2007. – 832 с.

2. Хастингс Н. Справочник по статистическим распределениям / Н. Хастингс, Дж. Пикок. – М.: Статистика, 1980. – 94 с.

3. Кирьянов Д. Самоучитель MATHCAD 2001 / Д.В. Кирьянов. – СПб.: БХВ – Петербург.– 2002. – 844 с.

4. Каплан А.В. Решение экономических задач на компьютере. / А.В. Каплан. – М.: Д МКПресс, СПб.: Петер, 2004. – 600 с.

5. Meeker W.Q. Statistical Methods for Reliability Data / W.Q. Meeker, L.A. Escobar. – New York.: John Wiley & Sons, inc., 1998. – 680 p.

6. Переверзев Е.С. Случайные сигналы в задачах оценки состояния технических систем / Е.С. Переверзев, Ю.Ф. Даниев, Г.П. Филей. – Киев.: Наук. думка, 1992. – 252 с.

7. Закс Л. Статистическое оценивание / Л. Закс.– М.: Статистика, 1976. – 597 с.

Поступила в редакцию 10.02.2009

Рецензент: д-р физ.-мат. наук, проф., заведующая кафедрой компьютерного моделирования и информационных технологий М.В. Новожилова, Харьковский государственный технический университет строительства и архитектуры, Харьков, Украина.

A COMPARATIVE ANALYSIS OF RANDOM NUMBER GENERATORS IN STATGRAPHICS AND MATHCAD SYSTEMS

V.J. Dubnitsky, A.G. Protsenko

A study in random number generators in STATGRAPHICS and MATHCAD systems. Any coincidence in distribution laws and their parameters shown to take place only for uniform distribution at any amount of data. For other random value distribution laws, namely: normal distribution, log normal distribution, gamma distribution, beta distribution, Cauchy distribution, Laplace distribution, exponential distribution, logistic distribution and Weibull distribution coincidence never occurred in all cases even at generating a thousand data sample.

Keywords: STATGRAPHICS, MATHCAD, random number generators, quality of random number generators, probability distributions.

ПОРІВНЯЛЬНИЙ АНАЛІЗ ДАТЧИКІВ ВИПАДКОВИХ ЧИСЕЛ СИСТЕМ STATGRAPHICS І MATHCAD

В.Ю. Дубницький, О.Г. Проценко

Виконано дослідження датчиків випадкових чисел систем STATGRAPHICS і MATHCAD. Показано, що при будь-якій кількості даних збіг законів розподілу та їх параметрів відмічений тільки для рівномірного розподілу. Для інших законів розподілу випадкових величин збіг відмічений не в усіх випадках навіть при генеруванні вибірки об'ємом тисячу даних.

Ключові слова: STATGRAPHICS, MATHCAD, датчики випадкових чисел, якість датчиків випадкових чисел, закони розподілу.

Дубницький Валерій Юрьевич – канд. техн. наук, ст. науч. сотрудник, доцент кафедры высшей математики Харьковского института банковского дела Университета банковского дела НБУ, Украина, e-mail: valeriy_dubn@mail.ru.

Проценко Алексей Геннадьевич – студент 3 курса кафедры компьютерных систем и сетей Национального аэрокосмического университета им. Н.Е. Жуковского «ХАИ», Харьков, Украина.