УДК 004:519.2:519.6
П 77

ГЕНЕРУВАННЯ ВИПАДКОВИХ ЧИСЕЛ З РОЗПОДІЛОМ ГАУССА НА ОСНОВІ ПЕРЕТВОРЕННЯ ДЖОНСОНА ІЗ СІМ'Ї S_U

С. Б. Приходько, д-р техн. наук, доц.

Національний університет кораблебудування, м. Миколаїв

Анотація. Удосконалено метод генерування гауссівських випадкових чисел на основі нормалізуючого перетворення Джонсона із сім'ї S_U, який на відміну від існуючих методів для генерування одного значення гауссівської випадкової величини потребує тільки одного значення випадкової величини з рівномірним розподілом.
Ключові слова: випадкові числа, розподіл Гаусса, перетворення Джонсона.

Аннотация. Усовершенствован метод генерирования гауссовских случайных чисел на основе нормализирующего преобразования Джонсона из семейства S_U, который в отличие от существующих методов для генерирования одного значения гауссовской случайной величины требует только одного значения случайной величины с равномерным распределением.
Ключевые слова: случайные числа, распределение Гаусса, преобразование Джонсона.

Abstract. The method of generation of Gaussian random numbers using Johnson S_U-transformation has been improved. Unlike other existing methods of generation of a Gaussian random variable, the transformation requires only one value of the random variable with the uniform distribution.
Keywords: random numbers, Gaussian distribution, Johnson transformation.

ПОСТАНОВКА ПРОБЛЕМИ

Для статистичного моделювання різноманітних випадкових процесів та подій дуже часто виникає потреба в значеннях випадкових чисел з розподілом Гаусса. Сьогодні для їх генерування існує багато методів, які базуються на перетворенні випадкових чисел з рівномірним розподілом у такі, що мають розподіл Гаусса [1, 2, 4–6]. Частина цих методів базується на центральній граничній теоремі і потребує для генерації одного значення гауссівської випадкової величини від 12 значень випадкових чисел з рівномірним розподілом. Інші використовують спеціальні перетворення (наприклад, Бокса–Мюллера) і потребують на створення певної кількості значень гауссівської випадкової величини приблизно в 1,25 разу більше значень випадкових чисел з рівномірним розподілом. Усе це призводить до проблеми зменшення фактичної кількості значень псевдовипадкових чисел з рівномірним розподілом у межах періоду генератора, яку можна використовувати до їх повторення [2].

АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Для вирішення цієї проблеми відомий метод на основі зворотної функції [6]. Але його використання для генерування випадкових чисел з розподілом Гаусса ускладнено тим, що не існує аналітичного виразу функції розподілу Гаусса, а це потребує відповідної її апроксимації. У роботі [2] для моделювання гауссівських випадкових величин запропоновано застосовувати нормалізуюче перетворення Джонсона із сім'ї S_B, яке не є бієктивним (взаємно однозначним). А це призводить до поганих результатів на границях або «хвостах» розподілу. Для покращення результатів генерування випадкових чисел з розподілом Гаусса може бути застосовано бієктивне нормалізуюче перетворення, яким є перетворення Джонсона із сім'ї S_U.

МЕТА РОБОТИ – удосконалення методу генерування гауссівських випадкових чисел на основі нормалізуючого перетворення Джонсона із сім'ї S_U.

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

У загальному випадку перетворення Джонсона має вигляд [2]

$$z = \gamma + \eta h(x, \varphi, \lambda); \quad \eta > 0; \quad -\infty < \gamma < \infty; \quad \lambda > 0; \quad -\infty < \varphi < \infty, \tag{1}$$

де z – нормально розподілена випадкова величина з математичним сподіванням нуль і дисперсією одиниця; x – випадкова величина з розподілом Джонсона; γ, η, φ, λ – параметри перетворення або розподілу Джонсона; h – функція з певної сім'ї:

$$h = \begin{cases} \ln(\tilde{x}), & x > \varphi, & \text{для сім'ї } S_L; \\ \ln[\tilde{x}/(1-\tilde{x})], & \varphi < x < \varphi + \lambda, & \text{для сім'ї } S_B; \\ \mathrm{Arsh}(\tilde{x}), & -\infty \le x \le +\infty, & \text{для сім'ї } S_U. \end{cases}$$

Тут $\tilde{x} = (x - \varphi)/\lambda$; $\mathrm{Arsh}(\tilde{x}) = \mathrm{h}\left(\tilde{x} + \sqrt{\tilde{x}^2 + 1}\right)$.

Оцінки параметрів для обраної сім'ї перетворення (1) можна знайти шляхом розв'язання наступної задачі математичного програмування [2]:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\theta} \left\{ \sum_{j=1}^{m} \left[y(x_j) - f(x_j, \theta) \right]^2 \right\}, \tag{2}$$

де θ – вектор невідомих параметрів, $\theta = \{\gamma, \eta, \varphi, \lambda\}$; x_j – значення x в середині j-го підінтервалу; $y(x_j)$ –

значення ординати гістограми для x_j; $f(x_j,\theta)$ – вираз функції щільності ймовірності Джонсона для x_j; m – кількість підінтервалів гістограми.

Оцінювання параметрів перетворення Джонсона за розв'язком задачі (2) можна здійснити у разі великої вибірки, коли є можливість побудувати гістограму. У разі як малої, так і великої вибірок відповідні оцінки параметрів можуть бути визначені за розв'язком задачі [2]:

$$\hat{\theta} = \arg\min_{\theta}\left\{A_z^2 + (\varepsilon_z - 3)^2 + \bar{z}^2 + \left(S_z^2 - 1\right)^2\right\}, \quad (3)$$

де $A_z = \dfrac{1}{nS_z^3}\sum_{i=1}^{n}(z_i - \bar{z})^3$; $\varepsilon_z = \dfrac{1}{nS_z^4}\sum_{i=1}^{n}(z_i - \bar{z})^4$; $\bar{z} = \dfrac{1}{n}\sum_{i=1}^{n} z_i$;

$S_z^2 = \dfrac{1}{n}\sum_{i=1}^{n}(z_i - \bar{z})^2$; z_i – i-ті значення величини z у вибірці довжиною n, $i\in[1, n]$, обчислюються за (1).

У виразі (3) враховане те, що в перетворенні Джонсона (1) z – це нормально розподілена випадкова величина з нульовим математичним сподіванням і одиничною дисперсією.

Неважко побачити, що для отримання значення випадкової величини з розподілом Гаусса за перетворенням (1) треба мати значення випадкової величини з розподілом Джонсона, яке в свою чергу потрібно одержувати за значенням випадкової величини з рівномірним розподілом. Тоді суть методу моделювання гауссівських випадкових величин на основі перетворення Джонсона, що запропоновано в [2], полягає у наступному. Спочатку генеруємо значення випадкової величини з рівномірним розподілом, яке перетворюємо у значення випадкової величини з розподілом Джонсона певної сім'ї, а потім за допомогою (1) отримуємо значення випадкової величини з розподілом Гаусса.

Для практичної реалізації методу, що запропоновано в [2], потрібно мати функцію, за допомогою якої значення випадкової величини з рівномірним розподілом можна перетворити у значення випадкової величини з розподілом Джонсона певної сім'ї. Як таку функцію у [2] запропоновано використовувати функцію синуса, яка дозволяє отримати значення випадкової величини x з розподілом Джонсона із сім'ї S_B за значення випадкової величини U з рівномірним розподілом:

$$x = \sin\frac{\pi U}{2},\ U \in [-1, 1]. \quad (4)$$

Використовуючи (4), значення гауссівської випадкової величини з математичним сподіванням нуль і дисперсією одиниця в [2] запропоновано визначати за перетворенням Джонсона (1) із сім'ї S_B за формулою

$$z = 0{,}3522056\ln\left[\tilde{x}/(1-\tilde{x})\right], \quad (5)$$

де $\tilde{x} = \left(\sin\dfrac{\pi U}{2} + 1{,}0017416\right)\Big/2{,}0034832$.

Але використання (5) для генерування гауссівських випадкових чисел дає задовільні результати в межах $m_z \pm 2\sigma_z$. Тут m_z – це математичне сподівання випадкової величини z, а σ_z – її середнє квадратичне відхилення. На границях інтервалу $m_z \pm 2{,}5\sigma_z$ результати генерування стають незадовільними: емпірична абсолютна частота величини z перевищує теоретичну настільки, що з довірчою ймовірністю 0,95 нульова гіпотеза про нормальність закону розподілу повинна бути відкинута. Такий результат можна пояснити тим, що перетворення Джонсона (1) із сім'ї S_B не є бієктивним і множини значень випадкових величин z та x не є ізоморфними.

Виправити таку ситуацію можна за рахунок застосування перетворення Джонсона (1) із сім'ї S_U, яке є бієктивним, а множини значень випадкових величин z та x є ізоморфними. Як функцію, яка дозволяє отримати значення випадкової величини x з розподілом Джонсона (1) із сім'ї S_U за значенням випадкової величини U з рівномірним розподілом, пропонується використати функцію

$$x = \mathrm{tg}(c_1U + c_3U^3),\ U \in [-1, 1]. \quad (6)$$

Тут c_1 і c_3 – постійні параметри, $c_1 = 0{,}55$ і $c_3 = 1{,}0$.

За (6) отримують вибірку значень випадкової величини x з практично нульовою асиметрією та ексцесом, який дорівнює 61,56. Ці значення асиметрії та ексцесу за аналітичною залежністю, що наведена в [3], відповідають розподілу Джонсона із сім'ї S_U. Оцінки параметрів перетворення Джонсона із сім'ї S_U знаходимо за розв'язком задачі (3). Вони є такими: $\gamma = 0$; $\eta = 0{,}532275$; $\varphi = 0$; $\lambda = 0{,}258099$.

Тоді значення гауссівської випадкової величини з математичним сподіванням нуль і дисперсією одиниця знаходимо за перетворенням Джонсона (1) із сім'ї S_U з наведеними вище параметрами за формулою

$$z = 0{,}532275\,\mathrm{Arsh}(3{,}874868x). \quad (7)$$

За (7) можна здійснювати генерування гауссівських випадкових чисел з вибірковим середнім нуль і вибірковою дисперсією одиниця за таким алгоритмом. Спочатку генерують значення випадкової величини U з рівномірним розподілом на інтервалі [–1, 1]. Далі з виразу (6) за значенням величини U знаходять значення величини x, яке підставляють у (7) та отримують значення випадкової величини з розподілом Гаусса.

Для перевірки працездатності перетворення (7) та наведеного алгоритму було виконано генерування 2000 гауссівських випадкових чисел. Для знаходження значень величини U використовувався відомий алгоритмічний генератор

$$U_i = \frac{\omega_i - 32768}{32768};$$

$$\omega_i = (25173\,\omega_{i-1} + 13849)\mathrm{mod}65536.$$

Тут ω_i – i-те значення випадкової величини ω з рівномірним розподілом, $\omega \in [0, 65536)$. Початкове значення ω при генеруванні z дорівнювало 2009.

Отримані значення z знаходилися на інтервалі [–3,099, 3,111]. Вибіркове середнє та вибіркова дисперсія для z склали –0,0125 та 1,0324 відповідно. Значення асиметрії та ексцесу для z дорівнювали 0,05145 і 2,83200 відповідно.

За критерієм Пірсона (χ^2) для вибірки значень z було перевірено гіпотезу про нормальність їх розподілу. Значення χ^2 склало 6,19. Критичне значення χ^2 для рівня значимості 0,05 та кількості ступенів вільності 8 дорівнює 15,5. Останнє вказує на те, що гіпотезу про нормальність розподілу для вибірки значень z за критерієм Пірсона потрібно прийняти з довірчою ймовірністю 0,95.

Отримані результати свідчать про працездатність удосконаленого методу генерування гауссівських випадкових чисел за рахунок застосування нормалізуючого перетворення Джонсона із сім'ї S_U.

ВИСНОВКИ

Удосконалено метод генерування гауссівських випадкових чисел за рахунок використання нормалізуючого перетворення Джонсона із сім'ї S_U, що дозволяє як максимально використовувати наявні можливості генераторів псевдовипадкових чисел з рівномірним розподілом за рахунок того, що для створення одного значення гауссівської випадкової величини необхідне тільки одне значення випадкової величини з рівномірним розподілом, так і покращити результати генерування біля границь емпіричного розподілу. У подальшому планується отримання нових формул для обчислення значення випадкової величини з розподілом Джонсона із сім'ї S_U за значенням випадкової величини з рівномірним розподілом.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] Вероятностные методы в вычислительной технике [Текст] : учеб. пособие для вузов по спец. ЭВМ / А. В. Крайников, Б. А. Курдиков, А. Н. Лебедев [и др.] ; под ред. А. Н. Лебедева и Е. А. Чернявского. – М. : Высш. шк., 1986. – 312 с.

[2] **Приходько, С. Б.** Моделювання гаусівських випадкових величин на основі перетворення Джонсона із сім'ї S_B [Текст] / С. Б. Приходько // Інформатика та математичні методи в моделюванні. – 2012. – Т. 2, № 1. – С. 64–69.

[3] **Приходько, С. Б.** Аналитическая зависимость для выбора распределения Джонсона семейства S_L [Текст] / С. Б. Приходько, Л. Н. Макарова // Вестник ХНТУ. – Херсон : ХНТУ, 2012. – № 2 (45). – С. 101–104.

[4] **Ситнік, В. Ф.** Імітаційне моделювання [Текст] : навч. посіб. / В. Ф. Ситнік, Н. С. Орленко. – К. : КНЕУ, 1998. – 232 с.

[5] **Форсайт, Дж.** Машинные методы математических вычислений [Текст] / Дж. Форсайт, М. Малькольм, К. Моулер. – М. : Мир, 1980. – 279 с.

[6] Gaussian Random Number Generators [Electronic resource] / D. B. Thomas, W. Luk, P. H. W. Leong, J. D. Villasenor // ACM Computing Surveys. – Vol. 39. – No. 4. – Article 11. – 2007. – P. 1–38. – Access mode: http://www.cse.cuhk.edu.hk.

Надійшла до редколегії 24.09.13
Статтю рекомендує до друку член редколегії Вісника НУК
д-р техн. наук, проф. *І. І. Коваленко*
Статтю розміщено у Віснику НУК № 4, 2013