Некоторые распределения, связанные с генерацией псевдослучайных чисел

Некруткин В.В., к.ф.-м.н., доцент кафедры статистического моделирования СПбГУ, vnekr@statmod.ru; Суровикина Т.О., студентка СПбГУ, tamara.surovikina@gmail.com

Аннотация

В работе изучается распределение $\mathrm{U}(S)$, согласованное с представлением чисел с плавающей точкой, и близкие к нему распределения.

Введение

Результатом работы генератора псевдослучайных чисел является последовательность чисел $u_1, \ldots, u_n, \ldots$, причем $u_i \in [0,1]$. Автоматически u_i оказываются представленными в памяти компьютера в формате чисел с плавающей точкой. Соответствующий стандарт (см. [1]) определяет решётку чисел, принадлежащих $[0,1]$ следующим образом. Фиксируются два целых числа B и S, задающие множество

$$X(S,B) = \{x_{jk}\} \cup \{0,1\} \subset [0,1]$$

так, что

$$x_{jk} = 2^{-j}\left(1 + k2^{-S}\right), \tag{1}$$

где $0 \le k \le 2^S - 1$, $1 \le j \le L = 2^{2^{B-1}-1}$.

Параметр B определяет количество бит, отведенных на хранение экспоненты, S — мантиссы. В дальнейшем нам будет удобнее рассматривать как характеристику решётки именно пару чисел (S,L), поэтому вместо $X(S,B)$ используется далее запись $X_{S,L}$. Одна из причин таких переобозначений заключается в удобстве интерпретации параметров S и L: L есть максимальное значение экспоненты числа, а S определяет длину мантиссы.

Естественно возникает вопрос о распределении, хорошо аппроксимирующем равномерное распределение и сосредоточенном на множестве $X_{S,L}$.

В [2] с этой целью рассматривается дискретное распределение $\mathrm{U}(S,L)$, задаваемое таблицей

$$\mathrm{U}(S,L): \quad \begin{pmatrix} x_{jk} \\ p_{jk}, \end{pmatrix}, \tag{2}$$

где $1 \le j \le L+1$, $0 \le k \le 2^S - 1$, значения x_{jk} заданы формулой (1), $p_{jk} = 2^{-S-j}$ при $1 \le j \le L$, $p_{L+1,0} = 2^{-L}$ и $p_{L+1,k} = 0$ для всех $k > 0$.

Особый интерес имеет предельный случай $L = \infty$. Обозначим соответствующее ему распределение как $\mathrm{U}(S)$, оно формально задается той же таблицей (2), только здесь $p_{jk} = 2^{-S-j}$ при всех $j \ge 1$.

Так как на практике используются огромные значения L, разница между распределениями $\mathrm{U}(S)$ и $\mathrm{U}(S, L)$ незначительна. Таким образом, анализ свойств $\mathrm{U}(S)$ (несомненно более простой) во многом можно распространить и на $\mathrm{U}(S, L)$.

В [2] распределение $\mathrm{U}(S)$ постулируется как естественное приближение равномерного распределения $\mathrm{U}(0, 1)$, сосредоточенное на множестве

$$X = X_S = \big\{\{x_{jk}\} \cup \{0\} \cup \{1\}\big\} \subset [0, 1] \tag{3}$$

с $x_{jk} = 2^{-j}\left(1 + k2^{-S}\right)$, где $k = 0, 1, \ldots, 2^S - 1$; $j \ge 1$. Такое распределение можно рассматривать как теоретическую модель, возникающую при построении генераторов псевдослучайных чисел, согласованных с плавающей точкой (см., например, [3] и [4]).

В настоящей работе исследуются свойства как распределения $\mathrm{U}(S)$, так и нескольких аналогичных распределений. Из-за недостатка места все доказательства опущены.s

Некоторые распределения с носителем X_S, аппроксимирующие $\mathrm{U}(0, 1)$

Оказывается, что $\mathrm{U}(S)$ получается как распределение «округленной вниз» случайной величины, имеющей равномерное распределение $\mathrm{U}(0, 1)$ на отрезке $[0, 1]$.

Исходя из стандарта IEEE 754-2008 (см. [1]) здесь мы будем использовать три способа проектирования чисел из $[0, 1]$ на множество X_S:

1. число x переходит в тот ближайший элемент множества X_S, который не превосходит x («округление вниз»);

2. число x переходит в тот ближайший элемент множества X_S, который не превосходит x («округление вверх»);

3. число x переходит в ближайший элемент множества X_S (используется по умолчанию).

В дальнейшем под $\lfloor x \rfloor$ будем понимать величину, получающуюся при округлении x «вниз», $\lceil x \rceil$ — при округлении «вверх», и $\lfloor x \rceil$ — при округлении до ближайшей точки из X_S.

Имеет место следующее утверждение.

Предложение 1 *Пусть* $\alpha \in \mathrm{U}(0,1)$ *и числа* x_{jk} *определены в* (1)*. Тогда*

1. $\mathcal{L}(\lfloor \alpha_S \rfloor) = \mathrm{U}(S)$*;*

2. Если $j \geq 1$ *и* $0 < k \leq 2^S$*, то* $\mathrm{P}(\lceil \alpha_S \rceil = x_{jk}) = 2^{-S-j}$*;*

3. Если $j \geq 1$ *и* $0 \leq k < 2^S$*, то*

$$\mathrm{P}(\lfloor \alpha_S \rceil = x_{jk}) = q_{jk} = \begin{cases} 2^{-S-j} & \text{при } j \geq 1 \text{ и } k \neq 0, \\ 3 \cdot 2^{-S-j-2} & \text{при } k = 0. \end{cases}$$

Кроме того, $\mathrm{P}(\lceil \alpha_S \rceil = 1) = 2^{-S-1}$ *и* $\mathrm{P}(\lfloor \alpha_S \rceil = 1) = 2^{-S-2}$.

В [2] отмечено, что, если случайные величины η и γ независимы, η равномерно распределена на множестве $\{0, \ldots, 2^S - 1\}$, а $\gamma \in \mathrm{Geom}(1/2)$, то случайная величина

$$\xi_S = 2^{-\gamma}\big(1 + \eta 2^{-S}\big) \tag{4}$$

имеет распределение $\mathrm{U}(S)$. В частности, мантисса и экспонента такой случайной величины являются независимыми.

Можно показать, что аналогичные результаты имеют место и для случайных величин $\lceil \alpha_S \rceil$ и $\lfloor \alpha_S \rceil$.

Предложение 2 *Пусть* η *и* γ *независимы, причем* $\gamma \in \mathrm{Geom}(1/2)$.

1. Если η *равномерно распределена на множестве* $\{1, \ldots, 2^S\}$*, то случайная величина* (4) *имеет распределение* $\mathcal{L}\big(\lceil \alpha_S \rceil\big)$.

2. Если η *— дискретная случайная величина с таблицей распределения*

$$\mathcal{L}(\eta): \begin{pmatrix} 0 & 1 & \ldots & 2^S - 1 & 2^S \\ 2^{-S-1} & 2^{-S} & \ldots & 2^{-S} & 2^{-S-1} \end{pmatrix},$$

то случайная величина (4) *имеет распределение* $\mathcal{L}\big(\lfloor \alpha_S \rceil\big)$.

Битовая структура распределения $\mathrm{U}(S)$

Рассмотрим случайную величину ξ_S, определенную в (4) и имеющую распределение $\mathrm{U}(S)$. Представим ее в виде

$$\xi_S = \sum_{i \geq 1} \beta_i 2^{-i},$$

где β_i — случайные величины, принимающие значения 0 и 1 (биты двоичного разложения ξ_S). Имеет место следующее утверждение.

Предложение 3 *1. Если* $i \geq 1$*, то*

$$\mathrm{P}(\beta_i = 1) = \begin{cases} 1/2 & \text{при } i \leq S+1, \\ 2^{-(i-S)} & \text{при } i > S+1. \end{cases} \qquad (5)$$

2. Пусть $1 < i < j$*. Тогда*

$$\mathrm{P}(\beta_i = 1, \beta_j = 1) = \begin{cases} 1/4 & \text{при } j \leq S, \\ 2^{-(j-S+1)} & \text{при } j > S \text{ и } j - i \leq S, \\ 0 & \text{при } j > S \text{ и } j - i > S. \end{cases} \qquad (6)$$

3. Биты $\beta_1, \ldots, \beta_{S+1}$ *независимы в совокупности.*

Следствие 1 *Из равенств* (5) *и* (6) *автоматически вытекает, что при* $i < j$ *случайные биты* β_i *и* β_j *являются независимыми тогда и только тогда, когда* $i \leq S$ *и* $j \leq i + S$*.*

Результат Предложения 3 может служить теоретической шкалой для сравнения генераторов псевдослучайных чисел между собой. В частности, вычислительные эксперименты показывают, что биты псевдослучайных чисел, порождаемые генератором SFMT [5], ведут себя в целом так же, как описано в этом предложении, в то время как распределения младших бит генератора WH [6] сильно отличаются от (5).

О распределении случайной величины $1 - \xi_S$

Хорошо известно, что, если $\alpha \in \mathrm{U}(0,1)$ то $1 - \alpha$ имеет такое же распределение. Так как $\mathrm{U}(S)$ рассматривается как некоторое дискретное приближение $\mathrm{U}(0,1)$, то возникает вопрос о том, сохраняется ли (хотя бы в пределе при $S \to \infty$) данное свойство для случайной величины ξ_S, имеющей распределение $\mathrm{U}(S)$.

Суть вопроса состоит в том, что мы имеем дело только с числами, лежащими на решетке X_S, определенной в (3), в то время как значения случайной величины $1 - \xi_S$, вообще говоря, не всегда принадлежат X_S.

Тем самым нас интересует «проекция» случайной величины $\xi_S^{(1)} \stackrel{\text{def}}{=} 1 - \xi_S$ на множество X_S, и мы снова приходим к трем способа округления, согласующимся со стандартом IEEE 754-2008.

Как и раньше, ξ_S принимает значения (1), где $j \geq 1, 0 \leq k \leq 2^S - 1$. Соответственно нас интересует решение уравнения вида $x_{jk} = 1 - x_{im}$, которое при $j > 1$ не всегда разрешимо. При $j = 1$ решение легко выписывается. Условия разрешимости для случая $j > 1$ описывает следующая лемма.

Лемма 1 *Пусть* $1 \le m \le 2^S - 1$ *и* $i \ge 1$. *Уравнение*

$$2^{-j}\bigl(1 + k2^{-S}\bigr) = 1 - 2^{-i}\bigl(1 + m2^{-S}\bigr) \qquad (7)$$

имеет решение (j, k) *с* $j > 1$ *тогда и только тогда, когда* $i = 1$ *и*

$$1 + 2^S - 2^{S-n} \le m \le 2^S - 2^{S-n-1}$$

для некоторого $0 \le n < S$.
При этом $j = n + 2$ *и* $k = 2^S - 2^{n+1}\bigl(m - 2^S + 2^{S-n}\bigr)$.

Используя Лемму 1, можно доказать следующие утверждения.

Предложение 4 *(Округление «вниз»)*
Пусть числа x_{jk} *при* $0 \le k < 2^S$, $j \ge 1$ *определены в* (1). *Тогда*

$$\mathrm{P}\left(\lfloor \xi_S^{(1)} \rfloor = x_{jk}\right) = q_{jk}, \quad \textit{где}$$

1. $q_{10} = 3 \cdot 2^{-S-2}$,
2. $q_{1k} = 2^{-S-1}$ *при* $k \ne 2^S\bigl(1 - 2^{-\ell}\bigr)$, $\ell = 1, \ldots, S-1$,
3. $q_{1k} = 2^{-S-\ell-2}\bigl(2^{\ell+1} + 1\bigr)$ *при* $k = 2^S\bigl(1 - 2^{-\ell}\bigr)$ *с* $\ell = 1, \ldots, S-1$,
4. $q_{1k} = 2^{-2S-1}\bigl(2^S + 1\bigr)$ *при* $k = 2^S - 1$.
5. *Пусть* $j > 1$. *Если пара* (j, k) *является решением уравнения* (7), *то* $q_{jk} = 2^{-S-1}$, *иначе* $q_{jk} = 0$.

Предложение 5 *(Округление до ближайшего значения).*
Пусть числа x_{jk} *при* $0 \le k \le 2^S$, $j \ge 1$ *определены в* (1). *Тогда*

$$\mathrm{P}\left(\lfloor \xi_S^{(1)} \rceil = x_{jk}\right) = q_{jk}, \quad \textit{где}$$

1.

$$q_{1k} = \begin{cases} 3 \cdot 2^{-S-2} & \textit{при } k = 0, \\ 2^{-S-1} & \textit{при } 0 < k < 2^S, \\ 2^{-S-2} & \textit{при } k = 2^S. \end{cases}$$

2. *Пусть* $j > 1$. *Если пара* (j, k) *является решением уравнения* (7), *то* $q_{jk} = 2^{-S-1}$, *иначе* $q_{jk} = 0$.

Предложение 6 *(Округление «вверх»).*
Пусть числа x_{jk} *при* $0 \le k \le 2^S$, $j \ge 1$ *определены в* (1). *Тогда*

$$\mathrm{P}\left(\lceil \xi_S^{(1)} \rceil = x_{jk}\right) = q_{jk}, \quad \textit{где}$$

1. $q_{1k} = 2^{-S-1}$ *при* $k = 0, \ldots, 2^S$;
2. *Пусть* $j > 1$. *Если пара* (j,k) *является решением уравнения* (7), *то* $q_{jk} = 2^{-S-1}$, *иначе* $q_{jk} = 0$.

Отметим, что во всех трех случаях получившиеся распределения имеют конечные носители, причем соответствующие (случайные) мантисса и порядок уже не являются независимыми.

Чтобы понять, насколько получившиеся распределения отличаются между собой и от исходного распределения $\mathrm{U}(S)$, сосчитаем попарные расстояния по вариации между этими четырьмя распределениями, считая, что их общим носителем является множество X_S. Используя теорему Шеффе (см. [7, с. 306]), получим следующий результат.

Предложение 7 *Выполняются следующие равенства:*

1. $\rho_{var}\Big(\mathrm{U}(S), \mathcal{L}\big(\lfloor \xi_S^{(1)} \rfloor\big)\Big) = \rho_{var}\Big(\mathrm{U}(S), \mathcal{L}\big(\lceil \xi_S^{(1)} \rceil\big)\Big) =$
$= \rho_{var}\Big(\mathrm{U}(S), \mathcal{L}\big(\lceil \xi_S^{(1)} \rceil\big)\Big) = 1/3 + 4^{-S}/6$;
2. $\rho_{var}\Big(\mathcal{L}\big(\lfloor \xi_S^{(1)} \rfloor\big), \mathcal{L}\big(\lfloor \xi_S^{(1)} \rceil\big)\Big) = \rho_{var}\Big(\mathcal{L}\big(\lfloor \xi_S^{(1)} \rceil\big), \mathcal{L}\big(\lceil \xi_S^{(1)} \rceil\big)\Big) = 2^{-S-2}$;
3. $\rho_{var}\Big(\mathcal{L}\big(\lfloor \xi_S^{(1)} \rfloor\big), \mathcal{L}\big(\lceil \xi_S^{(1)} \rceil\big)\Big) = 2^{-S-1}$.

Таким образом, даже при больших S расстояние по вариации между $\mathrm{U}(S)$ и распределением любого способа округления величины $1 - \xi_S$ не стремится к нулю.

В то же время распределения $\mathcal{L}(\lfloor \xi_S^{(1)} \rfloor)$, $\mathcal{L}(\lceil \xi_S^{(1)} \rceil)$ и $\mathcal{L}(\lfloor \xi_S^{(1)} \rceil)$ равноудалены по вариации от $\mathrm{U}(S)$ и при больших значениях параметра S практически совпадают.

Заключение

Таким образом, в настоящей работе приведены естественные способы получения распределения $\mathrm{U}(S)$ и близких к нему распределений, описаны

распределения случайной величины $1 - \xi_S$ при различных условиях ее адаптации к числовой решетке, согласованной с представлением чисел с плавающей точкой, и проанализирована степень их близости.

Кроме того, изучена битовая структура случайной величины, имеющей распределение $\mathrm{U}(S)$.

Литература

[1] IEEE Standard for Floating-Point Arithmetic, IEEE Std 754–2008. — P. 1–58.

[2] Nekrutkin V. On the complexity of binary floating point pseudorandom generation // Monte Carlo Methods and Applications. — 2016. — Vol. 22. — P. 109–116.

[3] Fog A. Pseudo random number generators. — URL: http://www.agner.org/random/.

[4] Morgenstern T. Uniform Random Rational Number Generation // Operations Research Proceedings 2006 / Ed. by Karl-Heinz Waldmann, Ulrike M. Stocker. — Springer. — P. 569–574.

[5] Matsumoto M., Saito M. SIMD-Oriented Fast Mersenne Twister: a 128-bit Pseudoran dom Number Generator // Monte Carlo and Quasi-Monte Carlo Methods 2006 / Ed. by A. Keller, S. Heinrich, H Niederreiter. — Springer, 2008. — P. 607–622.

[6] Wichmann B. A., Hill I. D. Algorithm AS 183: An Efficient and Portable Pseudo-Ran dom Number Generator // Journal of the Royal Statistical Society. Series C (Applied Statistics). — 1982. — Vol. 31. — P. 188–190.

[7] Биллингсли П. Сходимость вероятностных мер. — Наука, М., 1977.