



# Math-Net.Ru

Общероссийский математический портал

В. А. Ватутин, О близости распределения некоторой случайной величины к равновероятному распределению, *Матем. вопр. криптогр.*, 2023, том 14, выпуск 1, 5–14

DOI: 10.4213/mvk427

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением  
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 18.97.14.89

22 января 2025 г., 01:13:43



## О близости распределения некоторой случайной величины к равновероятному распределению

В. А. Ватутин

*Математический институт им В.А. Стеклова  
Российской академии наук, Москва*

*Получено 12.V.2022*

**Аннотация.** Пусть  $b \geq 2$  и  $N$  – натуральные числа,  $X_0, X_1, \dots, X_{n-1}$  – неоднородная последовательность независимых случайных величин, принимающих значения  $0, 1, \dots, b-1$ ,

$$Y_n = X_0 + bX_1 + \dots + X_{n-2}b^{n-2} + X_{n-1}b^{n-1}$$

и

$$Z_n = Y_n \bmod N.$$

В предположении, что числа  $b$  и  $N$  взаимно просты, оценивается близость распределения случайной величины  $Z_n$  к равновероятному распределению на множестве  $\{0, 1, \dots, N-1\}$ .

**Ключевые слова:** преобразования случайных величин, приведение по модулю, равновероятные распределения, расстояния между распределениями

## On the closeness of distribution of some random variable to the equiprobable one

V. A. Vatutin

*Steklov Mathematical Institute of Russian Academy of Sciences, Moscow*

**Abstract.** Let  $b \geq 2$  and  $N$  be natural numbers,  $X_0, X_1, \dots, X_{n-1}$  be nonhomogeneous sequence of independent random variables taking values  $0, 1, \dots, b-1$ ,

$$Y_n = X_0 + bX_1 + \dots + X_{n-2}b^{n-2} + X_{n-1}b^{n-1}$$

and

$$Z_n = Y_n \bmod N.$$

We estimate the closeness of distribution of random variable  $Z_n$  to the uniform distribution on  $\{0, 1, \dots, N-1\}$  in the case when  $b$  and  $N$  are mutually prime.

**Keywords:** transformations of random variables, modulo reduction, equiprobable distributions, distances between distributions

Пусть  $X_0, X_1, \dots, X_{n-1}$  – неоднородная последовательность независимых случайных величин со значениями в множестве  $\{0, 1, \dots, b-1\}$  и распределениями, задаваемыми соотношениями

$$\mathbf{P}(X_i = t) = p_{it}, \quad t = 0, 1, \dots, b-1, \quad i = 0, 1, \dots, n-1.$$

Свяжем с этой последовательностью неотрицательное целое число

$$Y_n = X_0 + X_1b + \dots + X_{n-2}b^{n-2} + X_{n-1}b^{n-1}. \quad (1)$$

Пусть далее  $N < b^n$  – натуральное число, а

$$Z_n = Y_n \bmod N. \quad (2)$$

Наша цель – изучить вероятностные свойства случайной величины  $Z_n$ , точнее, нас будет интересовать ответ на вопрос: при каких значениях чисел  $b$  и  $N$  распределение  $Z_n$  близко к равновероятному?

Задача, рассматриваемая в данной работе, тесно связана с проблемой моделирования знака  $\varkappa$ , имеющего равновероятное распределение на множестве  $\{0, 1, \dots, q-1\}$ , исходя из случайной последовательности  $\gamma = \{\gamma_n, n \geq 1\}$  независимых знаков, распределенных не обязательно равновероятно на множестве  $\{0, 1, \dots, p-1\}$  (см. в этой связи основополагающую работу фон Неймана [2], а также статью Ф. М. Малышева [1] и ссылки в ней). В анализируемой нами ситуации случайная величина  $Y_n$  (аналог величины  $\gamma_n$  из работы [1]) также имеет неравновероятное распределение, хотя некоторые знаки в наборе  $(X_0, X_1, \dots, X_{n-1})$  распределены равновероятно на множестве  $\{0, 1, \dots, b-1\}$ . Поскольку кроме преобразования (2) никаких других операций над случайной величиной  $Y_n$  производиться не будет, то в качестве основной характеристики распределения случайной величины  $Z_n$  будет служить параметр

$$\Delta_N(n) = \max_{0 \leq k \leq N-1} \left| \mathbf{P}(Z_n = k) - \frac{1}{N} \right|,$$

отражающий близость распределения  $Z_n$  к равновероятному распределению на множестве  $\{0, 1, \dots, N-1\}$ .

В данной работе мы получим оценки сверху величины  $\Delta_N(n)$ , накладывая на пару чисел  $(b, N)$  некоторые ограничения. А именно, будем

предполагать, что числа  $b$  и  $N$  взаимно просты и  $N < b^n$ . Обозначим через  $[x]$  целую часть числа  $x$  и запишем число  $H = [b^n/N]$  в виде

$$H = \left[ \frac{b^n}{N} \right] = rb^m + s, \tag{3}$$

где  $m$ ,  $r$  и  $s$  – целые числа, причем  $1 \leq r \leq b$ , а  $|s| < b^m/2$ . Отметим, что

$$rb^m + s \leq \frac{b^n}{N} < rb^m + s + 1$$

и, следовательно,

$$\left| \frac{1}{N} - \frac{r}{b^{n-m}} \right| \leq \frac{|s| + 1}{b^n}. \tag{4}$$

Для  $k = 0, 1, \dots, N - 1$  обозначим

$$\mathcal{A}_k := \{0 \leq j \leq H : k + jN \leq b^n - 1\}.$$

В силу определения величины  $H$  для любого  $j \in \mathcal{A}_k$  имеем

$$0 \leq j \leq \frac{b^n}{N} - \frac{k+1}{N} = H + \theta - \frac{k+1}{N},$$

где  $\theta \in [0, 1)$ . Таким образом, либо  $\mathcal{A}_k = \{0, 1, \dots, H - 1\}$ , либо  $\mathcal{A}_k = \{0, 1, \dots, H\}$ . Пусть, далее,

$$\mathcal{B}_k(r, s) := \mathcal{A}_k \setminus \{0, 1, \dots, rb^m - 1\},$$

если  $s$  в представлении (3) неотрицательно, и

$$\mathcal{C}_k(r, s) := \{0, 1, \dots, rb^m - 1\} \setminus \mathcal{A}_k,$$

если  $s$  в представлении (3) отрицательно. Обозначая  $|\mathcal{D}|$  мощность множества  $\mathcal{D}$ , имеем

$$|\mathcal{B}_k(r, s)| = |\mathcal{A}_k| - rb^m \leq H + 1 - rb^m \leq s + 1, \quad s > 0, \tag{5}$$

и

$$|\mathcal{C}_k(r, s)| = rb^m - |\mathcal{A}_k| \leq rb^m - H + 1 \leq 1 - s = 1 + |s|. \tag{6}$$

В формулируемой ниже лемме предполагаем, что случайные величины  $X_i, i \in [m, n - 1]$ , т.е. старшие разряды числа  $Y_n$  в его представлении (1) в  $b$ -ичной системе, имеют равномерное распределение на множестве  $\{0, 1, \dots, b - 1\}$ .

**Лемма 1.** Если случайные величины  $X_0, X_1, \dots, X_{n-1}$  независимы, причем

$$\mathbf{P}(X_i = t) = \frac{1}{b} \quad \text{для } t = 0, 1, \dots, b-1, \quad i = m, m+1, \dots, n-1, \quad (7)$$

то при  $s \geq 0$

$$\mathbf{P}(Z_n = k) = \frac{r}{b^{n-m}} + \frac{1}{b^{n-m}} \sum_{j \in B_k(r,s)} \mathbf{P}(Y_m = (k + Nj) \bmod b^m), \quad (8)$$

а при  $s < 0$

$$\mathbf{P}(Z_n = k) = \frac{r}{b^{n-m}} - \frac{1}{b^{n-m}} \sum_{j \in C_k(r,s)} \mathbf{P}(Y_m = (k + Nj) \bmod b^m). \quad (9)$$

*Доказательство.* Зафиксируем число  $k \in \{0, 1, \dots, N-1\}$  и для  $j \in \{0, 1, \dots, H\}$  положим

$$k_j := \left[ \frac{k + Nj}{b^m} \right].$$

Тогда для некоторого набора

$$(\zeta_0, \dots, \zeta_{m-1}) \in \{0, 1, \dots, b-1\}^m,$$

зависящего от  $k$  и  $j$ , справедливо представление

$$\begin{aligned} k + Nj &= (k + Nj) \bmod b^m + k_j b^m \\ &= \zeta_0 + \zeta_1 b + \dots + \zeta_{m-2} b^{m-2} + \zeta_{m-1} b^{m-1} + k_j b^m \\ &=: Y_m + k_j b^m. \end{aligned}$$

Отсюда следует, что  $b$ -ичное разложение числа  $k + jN \leq b^n - 1$  имеет вид

$$\begin{aligned} k + jN &= Y_m + k_j b^m \\ &= Y_m + \zeta_m b^m + \zeta_{m+1} b^{m+1} + \dots + \zeta_{n-2} b^{n-2} + \zeta_{n-1} b^{n-1}, \end{aligned}$$

где

$$(\zeta_0, \dots, \zeta_{m-1}, \zeta_m, \dots, \zeta_{n-1}) \in \{0, 1, \dots, b-1\}^n$$

– некоторый набор, зависящий от  $k$  и  $j$ .

Если  $j \in \mathcal{A}_k$ , то в силу условия (7)

$$\begin{aligned} \mathbf{P}(Y_n = k + jN) &= \mathbf{P}((X_0, X_1, \dots, X_{n-1}) = (\zeta_0, \zeta_1, \dots, \zeta_{n-1})) \\ &= \frac{1}{b^{n-m}} \mathbf{P}((X_0, X_1, \dots, X_{m-1}) = (\zeta_0, \zeta_1, \dots, \zeta_{m-1})) \\ &= \frac{1}{b^{n-m}} \mathbf{P}(Y_m = (k + Nj) \bmod b^m). \end{aligned} \quad (10)$$

Следовательно,

$$\begin{aligned} \mathbf{P}(Z_n = k) &= \sum_{j \in \mathcal{A}_k} \mathbf{P}(Y_n = k + jN) \\ &= \frac{1}{b^{n-m}} \sum_{j \in \mathcal{A}_k} \mathbf{P}(Y_m = (k + Nj) \bmod b^m). \end{aligned} \quad (11)$$

Рассмотрим сначала случай  $H = rb^m + s, s \geq 0$ . Тогда

$$\begin{aligned} \sum_{j \in \mathcal{A}_k} \mathbf{P}(Y_m = (k + Nj) \bmod b^m) &= \sum_{0 \leq j \leq rb^m - 1} \mathbf{P}(Y_m = (k + Nj) \bmod b^m) \\ &\quad + \sum_{j \in \mathcal{B}_k(r,s)} \mathbf{P}(Y_m = (k + Nj) \bmod b^m). \end{aligned} \quad (12)$$

Ясно, что для любого  $y \in \{0, 1, \dots, r - 1\}$

$$\begin{aligned} \sum_{yb^m \leq j \leq (y+1)b^m - 1} \mathbf{P}(Y_m = (k + Nj) \bmod b^m) \\ = \sum_{0 \leq q \leq b^m - 1} \mathbf{P}(Y_m = (k + Nq) \bmod b^m). \end{aligned}$$

Следовательно,

$$\begin{aligned} \sum_{0 \leq j \leq rb^m - 1} \mathbf{P}(Y_m = (k + Nj) \bmod b^m) \\ = \sum_{y=0}^{r-1} \sum_{yb^m \leq j \leq (y+1)b^m - 1} \mathbf{P}(Y_m = (k + Nj) \bmod b^m) \\ = r \sum_{0 \leq q \leq b^m - 1} \mathbf{P}(Y_m = (k + Nq) \bmod b^m). \end{aligned} \quad (13)$$

Поскольку числа  $N$  и  $b$  взаимно просты, то множество

$$\mathcal{D}_k := \{(k + Nq) \bmod b^m, q = 0, 1, \dots, b^m - 1\}$$

совпадает с множеством

$$\mathcal{D}_0 := \{(Nq) \bmod b^m, q = 0, 1, \dots, b^m - 1\} = \{0, 1, \dots, b^m - 1\}.$$

Поэтому для любого допустимого значения  $k$

$$\begin{aligned} \sum_{0 \leq q \leq b^m - 1} \mathbf{P}(Y_m = (k + Nq) \bmod b^m) \\ &= \sum_{0 \leq q \leq b^m - 1} \mathbf{P}(Y_m = Nq \bmod b^m) \\ &= \sum_{0 \leq q \leq b^m - 1} \mathbf{P}(Y_m = q) = 1. \end{aligned} \quad (14)$$

Объединяя соотношения (11)–(14), приходим к (8).

Случай  $H = rb^m + s, s < 0$ , исследуется таким же способом:

$$\begin{aligned} \mathbf{P}(Z_n = k) &= \frac{1}{b^{n-m}} \sum_{j \in \mathcal{A}_k} \mathbf{P}(Y_m = (k + Nj) \bmod b^m) \\ &= \frac{1}{b^{n-m}} \sum_{0 \leq j \leq rb^m - 1} \mathbf{P}(Y_m = (k + Nj) \bmod b^m) \\ &\quad - \frac{1}{b^{n-m}} \sum_{j \in \mathcal{C}_k(r,s)} \mathbf{P}(Y_m = (k + Nj) \bmod b^m) \\ &= \frac{r}{b^{n-m}} - \frac{1}{b^{n-m}} \sum_{j \in \mathcal{C}_k(r,s)} \mathbf{P}(Y_m = (k + Nj) \bmod b^m), \end{aligned} \quad (15)$$

что доказывает (9).

Лемма 1 доказана.  $\square$

**Замечание 1.** Легко проверить, что при доказательстве соотношений (14) независимость случайных величин  $X_0, \dots, X_{m-1}$  не использовалась. Таким образом, заключение леммы 1 остается верным в ситуации, когда набор как угодно зависимых между собой случайных величин  $X_0, \dots, X_{m-1}$  не зависит от набора независимых случайных величин  $X_m, \dots, X_{n-1}$ , каждая из которых имеет равномерное распределение на множестве  $\{0, 1, \dots, b - 1\}$ .

Пусть

$$h_i = \max_{0 \leq t \leq b-1} p_{it}, \quad i = 0, 1, \dots, m - 1.$$

**Теорема 1.** Если случайные величины  $X_0, X_1, \dots, X_{n-1}$  независимы, выполнено условие (7), числа  $b$  и  $N$  взаимно просты и  $H = rb^m + s$ , то

$$\Delta_N(n) = \max_{0 \leq k \leq N-1} \left| \mathbf{P}(Z_n = k) - \frac{1}{N} \right| \leq \frac{|s|+1}{b^n} + \frac{|s|+1}{b^{n-m}} \prod_{i=0}^{m-1} h_i.$$

*Доказательство.* Пусть  $s \geq 0$ . Используя соотношения (8) и обращаясь к (4) и (5), заключаем, что

$$\begin{aligned} \Delta_N(n) &= \max_{0 \leq k \leq N-1} \left| \mathbf{P}(Z_n = k) - \frac{1}{N} \right| \\ &\leq \left| \frac{r}{b^{n-m}} - \frac{1}{N} \right| + \max_{0 \leq k \leq N-1} \left| \mathbf{P}(Z_n = k) - \frac{r}{b^{n-m}} \right| \\ &\leq \frac{s+1}{b^n} + \frac{1}{b^{n-m}} \max_{0 \leq k \leq N-1} \left( \sum_{j \in \mathcal{B}_k(r,s)} \mathbf{P}(Y_m = (k + Nj) \bmod b^m) \right) \\ &\leq \frac{s+1}{b^n} + \frac{s+1}{b^{n-m}} \max_{0 \leq k \leq N-1, j \in \mathcal{B}_k(r,s)} \mathbf{P}(Y_m = (k + Nj) \bmod b^m) \\ &\leq \frac{s+1}{b^n} + \frac{s+1}{b^{n-m}} \prod_{i=0}^{m-1} h_i, \end{aligned} \tag{16}$$

что доказывает теорему для случая  $s \geq 0$ .

При  $s < 0$ , опираясь на (4), (6) и (9), имеем

$$\begin{aligned} \Delta_N(n) &= \max_{0 \leq k \leq N-1} \left| \mathbf{P}(Z_n = k) - \frac{1}{N} \right| \\ &\leq \left| \frac{r}{b^{n-m}} - \frac{1}{N} \right| + \max_{0 \leq k \leq N-1} \left| \mathbf{P}(Z_n = k) - \frac{r}{b^{n-m}} \right| \\ &\leq \frac{|s|+1}{b^n} + \frac{|s|}{b^{n-m}} \max_{0 \leq k \leq N-1, j \in \mathcal{C}_k(r,s)} \mathbf{P}(Y_m = (k + Nj) \bmod b^m) \\ &\leq \frac{|s|+1}{b^n} + \frac{|s|+1}{b^{n-m}} \prod_{i=0}^{m-1} h_i. \end{aligned}$$

Теорема 1 доказана. □

Конечно, условие

$$\mathbf{P}(X_i = t) = \frac{1}{b}, \quad t = 0, 1, \dots, b-1$$



при  $i = m, m+1, \dots, n-1$  является весьма ограничительным. Рассмотрим более общую ситуацию. С этой целью введем обозначение

$$g_i := \min_{0 \leq t \leq b-1} p_{it}, \quad i = 0, 1, \dots, n-1.$$

**Теорема 2.** Пусть случайные величины  $X_0, X_1, \dots, X_{n-1}$  независимы, числа  $b$  и  $N$  взаимно просты и  $H = rb^m + s$ . Предположим, что для некоторых чисел  $\varepsilon_1 \in (0, 1)$  и  $\varepsilon_2 > 0$  справедливы неравенства

$$\frac{1 - \varepsilon_1}{b^{n-m}} \leq \prod_{i=m}^{n-1} g_i \leq \prod_{i=m}^{n-1} h_i \leq \frac{1 + \varepsilon_2}{b^{n-m}} \quad (17)$$

и, кроме того, для некоторого  $\delta > 0$  справедлива оценка

$$\prod_{i=0}^{m-1} h_i \leq \frac{1 + \delta}{b^m}. \quad (18)$$

Тогда при  $s \geq 0$

$$\left| \mathbf{P}(Z_n = k) - \frac{1}{N} \right| \leq \frac{s+1}{b^n} + \frac{1}{b^{n-m}} \max \left\{ r\varepsilon_2 + \frac{(1+\varepsilon_2)(s+1)(1+\delta)}{b^m}, r\varepsilon_1 \right\},$$

а при  $s < 0$

$$\left| \mathbf{P}(Z_n = k) - \frac{1}{N} \right| \leq \frac{|s|+1}{b^n} + \frac{1}{b^{n-m}} \max \left\{ r\varepsilon_1 + \frac{(1+\varepsilon_1)|s|(1+\delta)}{b^m}, r\varepsilon_2 \right\}.$$

*Доказательство.* По аналогии с (10) и (11) при  $k + jN < b^n$  имеем

$$\begin{aligned} \mathbf{P}(Y_n = k + jN) &= \mathbf{P}((X_0, X_1, \dots, X_{n-1}) = (\zeta_0, \zeta_1, \dots, \zeta_{n-1})) \\ &= \mathbf{P}((X_m, \dots, X_{n-1}) = (\zeta_m, \dots, \zeta_{n-1})) \mathbf{P}(Y_m = (k + Nj) \pmod{b^m}) \\ &\leq \left( \prod_{i=m}^{n-1} h_i \right) \mathbf{P}(Y_m = (k + Nj) \pmod{b^m}) \\ &\leq \frac{1 + \varepsilon_2}{b^{n-m}} \mathbf{P}(Y_m = (k + Nj) \pmod{b^m}) \end{aligned}$$

и

$$\begin{aligned} \mathbf{P}(Y_n = k + jN) &\geq \left( \prod_{i=m}^{n-1} g_i \right) \mathbf{P}(Y_m = (k + Nj) \pmod{b^m}) \\ &\geq \frac{1 - \varepsilon_1}{b^{n-m}} \mathbf{P}(Y_m = (k + Nj) \pmod{b^m}). \end{aligned}$$

Отсюда следует, что при  $s \geq 0$

$$\begin{aligned} \mathbf{P}(Z_n = k) &= \sum_{0 \leq j \leq rb^{m-1}} \mathbf{P}(Y_n = k + jN) \\ &+ \sum_{j \in \mathcal{B}_k(r,s)} \mathbf{P}(Y_n = (k + Nj) \pmod{b^m}) \\ &\leq \frac{(1 + \varepsilon_2)}{b^{n-m}} \sum_{0 \leq j \leq rb^{m-1}} \mathbf{P}(Y_m = k + jN) \\ &+ \frac{(1 + \varepsilon_2)}{b^{n-m}} \sum_{j \in \mathcal{B}_k(r,s)} \mathbf{P}(Y_m = (k + Nj) \pmod{b^m}) \\ &\leq \frac{(1 + \varepsilon_2)}{b^{n-m}} \left( r + (s + 1) \prod_{i=0}^{m-1} h_i \right) \\ &\leq \frac{(1 + \varepsilon_2)}{b^{n-m}} \left( r + \frac{(s + 1)(1 + \delta)}{b^m} \right) \end{aligned}$$

и

$$\begin{aligned} \mathbf{P}(Z_n = k) &\geq \sum_{0 \leq j \leq rb^{m-1}} \mathbf{P}(Y_n = k + jN) \\ &\geq \frac{(1 - \varepsilon_1)}{b^{n-m}} \sum_{0 \leq j \leq rb^{m-1}} \mathbf{P}(Y_m = k + jN) \\ &= \frac{r(1 - \varepsilon_1)}{b^{n-m}}. \end{aligned}$$

Эти неравенства в сочетании с оценкой (4) влекут первое утверждение теоремы 2.

При  $s < 0$  имеем

$$\begin{aligned} \mathbf{P}(Z_n = k) &= \sum_{0 \leq j \leq \mathcal{A}_k} \mathbf{P}(Y_n = k + jN) \\ &\leq \sum_{0 \leq j \leq rb^m - 1} \mathbf{P}(Y_n = k + jN) \\ &\leq \prod_{i=m}^{n-1} h_i \sum_{0 \leq j \leq rb^m - 1} \mathbf{P}(Y_m = k + jN) \leq \frac{(1 + \varepsilon_2) r}{b^{n-m}} \end{aligned}$$

и

$$\begin{aligned} \mathbf{P}(Z_n = k) &= \sum_{0 \leq j \leq rb^m - 1} \mathbf{P}(Y_n = k + jN) \\ &\quad - \sum_{j \in \mathcal{C}_k(r, s)} \mathbf{P}(Y_n = (k + Nj) \pmod{b^m}) \\ &\geq \prod_{i=m}^{n-1} g_i \sum_{0 \leq j \leq rb^m - 1} \mathbf{P}(Y_m = k + jN) \\ &\quad - \prod_{i=m}^{n-1} h_i \sum_{j \in \mathcal{C}_k(r, s)} \mathbf{P}(Y_m = (k + Nj) \pmod{b^m}) \\ &\geq \frac{(1 - \varepsilon_1)}{b^{n-m}} r - \frac{1 + \varepsilon_2}{b^{n-m}} |s| \prod_{i=0}^{m-1} h_i \\ &\geq \frac{(1 - \varepsilon_1)}{b^{n-m}} r - \frac{|s| (1 + \varepsilon_2) (1 + \delta)}{b^{n-m}}. \end{aligned}$$

Эти неравенства и оценка (4) подтверждают справедливость второго утверждения теоремы 2. □

**Замечание 2.** Из вида оценок, установленных в теоремах 1 и 2, следует, что распределение случайной величины  $Z_n$  будет наиболее близко к равновероятному, если в представлении (3) параметр  $|s|$  будет мал.

## Список литературы

- [1] Малышев Ф. М., “Моделирование равномерного распределения, устойчивое к неравновероятности исходных знаков”, *Дискретная математика*, **17**:4 (2005), 72–80.
- [2] Neumann J. von, “Various techniques used in connection with random digits”, *John von Neumann, Collected Works*, **V**, MacMillan, New York, 1963, 768–770.