ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

УДК 519.214.5
doi: 10.18101/2304-5728-2017-3-78-88

О РАСПРЕДЕЛЕНИИ ЧИСЕЛ ПАР И ТРОЕК ОДИНАКОВЫХ ЗНАКОВ НА ЦИКЛЕ ДВОИЧНОЙ МУЛЬТИЦИКЛИЧЕСКОЙ ПОСЛЕДОВАТЕЛЬНОСТИ

© Меженная Наталья Михайловна
кандидат физико-математических наук, доцент,
Московский государственный технический университет им. Н. Э. Баумана
Россия, 105005, г. Москва, ул. 2-я Бауманская, 5, стр. 1
E-mail: natalia.mezhennaya@gmail.com

В работе рассмотрено предельное распределение чисел пар и троек одинаковых знаков на цикле двоичной мультициклической случайной последовательности с r регистрами, заполненными независимыми в совокупности двоичными случайными величинами с равномерными распределениями, в двух случаях: 1) когда число регистров r остается фиксированным, а их длины стремятся к бесконечности; 2) когда r также стремится к бесконечности. В первом случае с помощью неравенства Берри — Эссеена получены равномерные оценки скорости сходимости.

Ключевые слова: мультициклическая последовательность; нормальная предельная теорема; число пар одинаковых знаков; число троек одинаковых знаков; неравенство Берри — Эссеена.

Введение

Пусть $\mathbf{X}_j = (X_0^{(j)}, \ldots, X_{n_j-1}^{(j)})$ — независимые между собой наборы по n_j случайных величин, $j = 1, \ldots, r$. Мультициклический генератор (см. [1]) вырабатывает псевдослучайную последовательность по правилу

$$Z_t = \left(\sum_{j=1}^{r} X_{t \bmod n_j}^{(j)} \right) \bmod 2, \quad t = 0, 1, \ldots. \tag{1}$$

Величину $T = n_1 \ldots n_r$ принято называть длиной цикла (1).

Мультициклический генератор позволяет получить псевдослучайную последовательность с гарантированно длинным периодом T при наличии достаточно большого числа регистров с нетождественными заполнениями. С появлением программных средств генерации псевдослучайных чисел интерес к нему существенно снизился. Но недавно выяснилось, что удается эффективно использовать современные методы теории вероятностей для исследования статистических свойств на всем цикле выходной последовательности мультициклического генератора. Это свойство выгодно отличает его от большинства других генераторов и обусловлено его специфической структурой. Свойства двоичной мультициклической случайной последовательности изучались в работах [2], [3], [4]. В [2] получено распределение числа единиц ξ_r на цикле мультициклической последовательности вида (1) при независимых в совокупности и равномерно распределенных знаках в регистрах. Получены условия сходимости распределения ξ_r к распределению произведения независимых стандартных нормальных случайных величин, когда число регистров r фиксировано, а их длины $n_1,\ldots,n_r \to \infty$. Также получены условия сходимости распределения ξ_r к логнормальному закону, когда $r \to \infty$ и $n_1,\ldots,n_r \to \infty$. В обоих случаях найдены оценки скорости сходимости. В работах [3] и [4] проведено исследование устойчивости предельных распределений, полученных в [2], при нарушении предположений о распределении знаков в регистрах генератора: в [3] рассмотрено распределение знаков, отличное от равномерного, а в [4] — зависимые заполнения внутри регистров.

Пусть

$$\eta_k = \sum_{0 \le i_1 < i_2 < i_k \le T-1} I\{Z_{i_1} = Z_{i_2} = \ldots = Z_{i_k}\} \tag{2}$$

число наборов по $k \ge 2$ одинаковых знаков на цикле мультициклической последовательности вида (1).

Задача о распределении η_k связана с задачей повторениях знаков или цепочек в последовательности независимых испытаний (см. [5], [6], [7]). В частности, в [5] получена пуассоновская аппроксимация числа пар совпавших цепочек в последовательности независимых случайных величин с полиномиальным распределением, а в [6] — обобщение этих результатов для кратных совпадений цепочек. В [7] изучено также число серий кратных повторений и момента первого появления кратного повторения цепочек. Свойства аналогичных статистик в цепи Маркова изучены в работах [8], [9], [10].

Предельное распределение числа пар и троек одинаковых знаков

Свойства распределений случайных величин η_k могут быть получены из описанных выше результатов работы [2]. Если обозначить через ξ_r — число единиц в $Z_0,\ldots,Z_{T-1}$, то из (2)

$$\eta_k = \sum_{0 \le i_1 < i_2 < i_k \le T-1} I\{Z_{i_1} = Z_{i_2} = \ldots = Z_{i_k} = 1\} +$$

$$+ \sum_{0 \le i_1 < i_2 < i_k \le T-1} I\{Z_{i_1} = Z_{i_2} = \ldots = Z_{i_k} = 0\} = C_{\xi_r}^k + C_{T-\xi_r}^k . \tag{3}$$

Сформулируем следующие условия:

A) число регистров $r \ge 2$, длины регистров $n_1,\ldots,n_r$ нечетны и взаимно просты;

Б) $\mathbf{X}_j = (X_0^{(j)}, \ldots, X_{n_j-1}^{(j)}), j = 1, \ldots, r,$ — независимые между собой случайные векторы длин n_j, образованные независимыми (в совокупности) двоичными случайными величинами с равномерным распределением, т. е.

$$\mathbf{P}\{X_k^{(j)} = 1\} = 1 - \mathbf{P}\{X_k^{(j)} = 0\} = \frac{1}{2}, \ k = 0, \ldots, n_j - 1, \ j = 1, \ldots, r.$$

Заметим, что при выполнении условия Б распределения случайных величин ξ_r и $T - \xi_r$, а значит и $C_{\xi_r}^k$ и $C_{T-\xi_r}^k$, совпадают. Двумерный закон распределения $C_{\xi_r}^k$ и $C_{T-\xi_r}^k$ сосредоточен на линии $\{(x, y) : x = C_n^k, y = C_{T-n}^k, \ n = 0, 1, \ldots, T-1\}$.

Утверждение 1. *Пусть выполнены условия А и Б. Тогда*

$$\mathbf{E}\eta_k = 2^{-k+1} C_T^k, \ k = 2, 3, \tag{4}$$

$$\mathbf{E}\eta_4 = \frac{1}{192}\left(T \prod_{j=1}^{r} (3n_j - 2) + (8T - 6T^2 + T^3 - 4) \right), \tag{5}$$

$$\mathbf{D}\eta_2 = \frac{1}{16} T \left(\prod_{j=1}^{r} (3n_j - 2) - T \right) < \frac{3^r T^2}{16}, \tag{6}$$

$$\mathbf{D}\eta_3 = \frac{1}{4}(T-2)^2 \mathbf{D}\eta_2 = \frac{1}{64} T (T-2)^2 \left(\prod_{j=1}^{r} (3n_j - 2) - T \right) < \frac{3^r T^4}{64}. \tag{7}$$

Замечание 1. Формулу для дисперсии η_4 не приводим ввиду ее громоздкости, однако ее несложно получить через центральные моменты биномиального закона.

Замечание 2. Из формул (4) и (5) видно, что свойства случайных величин η_2, η_3 принципиально отличаются от свойств случайных величин $\eta_4, \eta_5, \ldots$ Например, средние значения η_2, η_3 зависят только от длины цикла T, а $\mathbf{E}\eta_4$ уже зависит от длин отдельных регистров.

Обозначим через $d(F, G)$ — расстояние в равномерной метрике между функциями распределения F и G на действительной прямой $\mathbb{R}$:

$$d(F, G) = \sup_{x \in \mathbb{R}} |F(x) - G(x)|.$$

Определим случайные величины $\tilde{\eta}_2 = T^{-1}(4\eta_2 - T^2 + 2T)$ и $\tilde{\eta}_3 = \dfrac{2\eta_3}{T(T-2)} - \dfrac{T-4}{12}$, а также введем обозначения Φ_r для функции распределения произведения r независимых стандартных нормальных случайных величин, Ψ_r для функции распределения произведения r независимых случайных величин, распределенных по закону $\chi^2(1)$. Для краткости будем писать $\Phi_1 = \Phi, \ \Psi_1 = \Psi$.

Замечание 3. Случайные величины $\tilde{\eta}_2$ и $\tilde{\eta}_3$ равны между собой. Это будет видно из доказательства утверждения 2.

Утверждение 2. *Пусть выполнены условия А и Б. Тогда при* $k = 2,3$

$$d(F_{\tilde{\eta}_k}, \Psi_r) \le 2C_{BE} \sum_{i=1}^{r} n_i^{-1/2}. \tag{8}$$

Здесь $C_{BE} < 0,48$ *— константа из неравенства Берри — Эссеена для распределения суммы независимых одинаково распределенных случайных величин* (*см.* [11]).

Замечание 4. Из формулы (8) следует, что если число $r \ge 2$ фиксировано, а $n_1, \ldots, n_r \to \infty$, то $F_{\tilde{\eta}_k}(x) \to \Psi_r(x)$ при всех $x \in \mathbb{R}$.

Замечание 5. Плотность для функции распределения $\Psi_r(x)$ получена в работе [12] (формула (33)) и имеет вид

$$\psi_r(z) = \left(2\sqrt{\pi}\right)^{-r} G_{0r}^{r0}\left(2^{-r} z \,|\, -\frac{1}{2}\right),$$

где G_{ab}^{cd} — G-функция Мейера (см., например, [13], раздел 5.3, с. 203).

Замечание 6. Утверждение 2 можно переформулировать следующим образом: формула (8) эквивалентна тому, что при $k = 2,3$

$$d(F_{\ln \tilde{\eta}_k}, F_r) \le 2C_{BE} \sum_{i=1}^{r} n_i^{-1/2}, \tag{9}$$

где F_r — это функция распределения суммы r независимых случайных величин, имеющих плотность распределения $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{x/2} e^{-e^x/2}$, $x \in \mathbb{R}$.

Действительно, пусть ς имеет распределение $\chi^2(1)$. Тогда плотность распределения $\ln\varsigma$ равна

$$f_{\ln\varsigma}(x) = y'_x \frac{1}{\sqrt{2\pi y}} e^{-y/2}\Bigg|_{y=e^x} = \frac{1}{\sqrt{2\pi}} e^{x/2} e^{-e^x/2}.$$

Отсюда получаем (9).

Замечание 7. Функция распределения F_r может быть представлена как функция распределения суммы вида $2(|\nu_1| + \ldots + |\nu_r|)$, в которой случайные величины $\nu_1, \ldots, \nu_r$ независимы и одинаково распределены по стандартному нормальному закону. Такой функции распределения соответствует характеристическая функция вида

$$\frac{1}{(2\pi)^{r/2}}\left(e^{-\frac{(2t)^2}{2}}\left(1 + Erf\left(\frac{2it}{\sqrt{2}}\right)\right)\right)^r, \tag{10}$$

где через $Erf(z)$ обозначена стандартная функция ошибок: $Erf(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-y^2} dy$. Как нетрудно убедиться, характеристическая функция одного слагаемого $|\nu_j|$ равна $\frac{1}{(2\pi)^{1/2}} e^{-\frac{t^2}{2}} \left(1 + Erf\left(\frac{\mathrm{i}t}{\sqrt{2}}\right)\right)$. Отсюда сразу получаем характеристическую функцию (10).

Теперь перейдем к случаю, когда число регистров r также стремится к бесконечности. Здесь мы применим тот же подход, что и при доказательстве теоремы 4 работы [2].

Введем обозначения

$$a = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} \ln x \, e^{-x^2/2} dx = -\frac{1}{2}\left(\gamma + \ln 2\right) \approx -0,635,$$

$$\sigma^2 = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} (\ln x - a)^2 \, e^{-x^2/2} dx = \frac{\pi^2}{8} \approx 1,234$$

($\gamma \approx 0.577$ — константа Эйлера). Параметры a и σ^2 — это математическое ожидание и дисперсия логарифма модуля стандартной нормальной случайной величины.

Рассмотрим центрированную и нормированную случайную величину $\vartheta = \frac{\ln \tilde{\eta}_k - \mathbf{E} \ln \tilde{\eta}_k}{2\sigma\sqrt{r}}$, $k = 2,3$.

Утверждение 3. *Пусть* $r, n_1, \ldots, n_r \to \infty$, *выполнены условия А и Б. Тогда при всех* $x \in \mathbb{R}$

$$F_\vartheta(x) \to \Phi(x). \tag{11}$$

Замечание 8. Если $r, n_1, \ldots, n_r \to \infty$, выполнены условия А и Б и $\frac{1}{\sqrt{r}} \sum_{j=1}^{r} \frac{\ln n_j}{n_j} \to 0$, то $\mathbf{E} \ln \tilde{\eta}_2 = ar + o(r^{1/2})$. Значит, распределение случайной величины ϑ совпадает с распределением $\vartheta' = \frac{\ln \tilde{\eta}_2 - 2ra}{2\sigma\sqrt{r}}$ и (11) можно переписать в виде $F_{\vartheta'}(x) \to \Phi(x)$.

Замечание 9. В силу связи нормального и логарифмически нормального законов, (11) эквивалентна тому, что при $k = 2,3$ и всех $y > 0$

$$\mathbf{P}\{\tilde{\eta}_k < e^{\mathbf{E} \ln \tilde{\eta}_k} y^{2\sigma\sqrt{r}}\} \to \frac{1}{\sqrt{2\pi}} \int_0^y u^{-1} e^{-\ln^2 u/2} du$$

(в правой части стоит функция распределения логарифмически нормального закона с параметрами 0 и 1).

Доказательства утверждений

Доказательство утверждения 1. В работе [2] (формулы (4) и (5)) показано, что

$$\mathbf{E}\xi_r = \frac{T}{2},\ \ \mathbf{D}\xi_r = \frac{T}{4},\ \ \mathbf{E}\left(\xi_r - \frac{T}{2}\right)^4 = \frac{1}{16}T\prod_{j=1}^{r}(3n_j - 2). \tag{12}$$

Для вычисления математических ожиданий η_k, $k = 2,3,4$, преобразуем формулу (3):

$$\eta_2 = \frac{1}{2}\big(\xi_r(\xi_r - 1) + (T - \xi_r)(T - \xi_r - 1)\big) = \left(\xi_r - \frac{T}{2}\right)^2 + \frac{T^2}{4} - \frac{T}{2}, \tag{13}$$

$$\eta_3 = \frac{1}{6}\big(\xi_r(\xi_r - 1)(\xi_r - 2) + (T - \xi_r)(T - \xi_r - 1)(T - \xi_r - 2)\big) =$$

$$= \frac{1}{2}(T-2)\left[\left(\xi_r - \frac{T}{2}\right)^2 + \frac{T^2}{12} - \frac{T}{3}\right], \tag{14}$$

$$\eta_4 = \frac{1}{24}\big(\xi_r(\xi_r - 1)(\xi_r - 2)(\xi_r - 3) + (T - \xi_r)(T - \xi_r - 1)(T - \xi_r - 2)(T - \xi_r - 3)\big) =$$

$$= \frac{1}{12}\left(\xi_r - \frac{T}{2}\right)^4 + \frac{1}{24}\left(\xi_r - \frac{T}{2}\right)^2\left(22 - 18T + 3T^2\right) + \frac{T(T-2)(T-4)(T-6)}{192}. \tag{15}$$

Следовательно, из (12) – (15) получаем, что

$$\mathbf{E}\eta_2 = \mathbf{D}\xi_r + \frac{T^2}{4} - \frac{T}{2} = \frac{T}{4} + \frac{T^2}{4} - \frac{T}{2} = \frac{T(T-1)}{4}.$$

$$\mathbf{E}\eta_3 = \frac{1}{2}(T-2)\left(\mathbf{D}\xi_r + \frac{T^2}{12} - \frac{T}{3}\right) = \frac{1}{2}(T-2)\left(\frac{T}{4} + \frac{T^2}{12} - \frac{T}{3}\right) = \frac{T(T-1)(T-2)}{24},$$

$$\mathbf{E}\eta_4 = \frac{1}{12}\mathbf{E}\left(\xi_r - \frac{T}{2}\right)^4 + \frac{1}{24}\mathbf{D}\xi_r\left(22 - 18T + 3T^2\right) + \frac{T(T-2)(T-4)(T-6)}{192} =$$

$$= \frac{1}{192}T\prod_{j=1}^{r}(3n_j - 2) + \frac{1}{96}T\left(22 - 18T + 3T^2\right) + \frac{T(T-2)(T-4)(T-6)}{192} =$$

$$= \frac{1}{192}T\prod_{j=1}^{r}(3n_j - 2) + \frac{1}{192}(8T - 6T^2 + T^3 - 4).$$

Таким образом, формулы (4) и (5) доказаны. Аналогично, можно вычислить дисперсии η_2 и η_3 (см. (6) и (7))

$$\mathbf{D}\eta_2 = \mathbf{E}\left(\left(\xi_r - \frac{T}{2}\right)^2 + \frac{T^2}{4} - \frac{T}{2} - \frac{T^2 - T}{4}\right)^2 = \mathbf{E}\left(\left(\xi_r - \frac{T}{2}\right)^2 - \frac{T}{4}\right)^2 =$$

$$= \mathbf{E}\left(\xi_r - \frac{T}{2}\right)^4 - \frac{T}{2}\mathbf{D}\xi_r + \frac{T^2}{16} =$$
$$= \frac{1}{16}T\prod_{j=1}^{r}(3n_j - 2) - \frac{T^2}{8} + \frac{T^2}{16} = \frac{1}{16}T\left(\prod_{j=1}^{r}(3n_j - 2) - T\right) < \frac{3^r T^2}{16},$$

$$\mathbf{D}\eta_3 = \mathbf{E}\left(\frac{1}{2}(T-2)\left[\left(\xi_r - \frac{T}{2}\right)^2 + \frac{T^2}{12} - \frac{T}{3}\right] - \frac{T(T-1)(T-2)}{24}\right)^2 =$$
$$= \frac{1}{4}(T-2)^2\mathbf{E}\left(\left(\xi_r - \frac{T}{2}\right)^2 + \frac{T^2}{12} - \frac{T}{3} - \frac{T(T-1)}{12}\right)^2 =$$
$$= \frac{1}{4}(T-2)^2\mathbf{E}\left(\left(\xi_r - \frac{T}{2}\right)^2 - \frac{T}{4}\right)^2 = \frac{1}{4}(T-2)^2\mathbf{D}\eta_2.$$

Утверждение 1 доказано.

Доказательство утверждения 2. В теореме 2 работы [2] было показано, что при выполнении условий А и Б расстояние в равномерной метрике между функцией распределения $F_{\tilde{\xi}_r}$ случайной величины $\tilde{\xi}_r = T^{-1/2}(2\xi_r - T)$ и функцией распределения Φ_r оценивается сверху выражением

$$d(F_{\tilde{\xi}_r}, \Phi_r) \le C_{BE}\sum_{i=1}^{r} n_i^{-1/2}. \tag{16}$$

Из формулы (16) следует, что если число $r \ge 2$ фиксировано, а $n_1, \ldots, n_r \to \infty$, то $F_{\tilde{\xi}_r}(x) \to \Phi_r(x)$ при всех $x \in \mathbb{R}$.

Далее воспользуемся этим результатом и формулами (13) и (14), из которых следует, что

$$\eta_2 = \frac{1}{4}\tilde{\xi}_r^2 T + \frac{T^2}{4} - \frac{T}{2},$$
$$\eta_3 = \frac{1}{2}(T-2)\left(\tilde{\xi}_r^2 T + \frac{T^2}{12} - \frac{T}{3}\right).$$

Значит, случайные величины $\tilde{\eta}_2$ и $\tilde{\eta}_3$ равны между собой и имеют такой же закон распределения, как и $\tilde{\xi}_r^2$. Так как при $x \ge 0$

$$F_{\tilde{\xi}_r^2}(x) = F_{\tilde{\xi}_r}(\sqrt{x}) - F_{\tilde{\xi}_r}(-\sqrt{x}) = 2F_{\tilde{\xi}_r}(\sqrt{x}) - 1,$$
$$\Psi_r(x) = \Phi_r(\sqrt{x}) - \Phi_r(-\sqrt{x}) = 2\Phi_r(\sqrt{x}) - 1$$

(в силу симметричности распределений случайной величины $\tilde{\xi}_r$ и стандартного нормального), то

$$d(F_{\tilde{\xi}_r^2}, \Psi_r) \le \sup_{x \ge 0}\left|F_{\tilde{\xi}_r^2}(x) - \Psi_r(x)\right| = \sup_{x \ge 0}\left|2F_{\tilde{\xi}_r}(\sqrt{x}) - 2\Phi_r(\sqrt{x})\right| \le$$

$$\le 2d(F_{\tilde{\xi}_r},\Phi_r)\le 2C_{BE}\sum_{i=1}^{r} n_i^{-1/2}.$$

Так как распределения пар случайных величин $\tilde{\xi}_r^2$ и $\tilde{\eta}_2$, $\tilde{\xi}_r^2$ и $\tilde{\eta}_3$ совпадают, то из последней оценки получаем (8). Утверждение 2 доказано.

Доказательство утверждения 3. Заметим, что

$$\vartheta=\frac{\ln\tilde{\eta}_k-\mathbf{E}\ln\tilde{\eta}_k}{\sigma\sqrt{r}}=\frac{2\ln|\tilde{\xi}_r|-2\mathbf{E}\ln|\tilde{\xi}_r|}{\sqrt{4\mathbf{D}\ln|\tilde{\xi}_r|}}=\frac{\ln|\tilde{\xi}_r|-\mathbf{E}\ln|\tilde{\xi}_r|}{\sqrt{\mathbf{D}\ln|\tilde{\xi}_r|}}. \tag{17}$$

Обозначим S_j — число единиц в j-м регистре, $j=1,\ldots,r.$ Из леммы 1 работы [2] имеем

$$\tilde{\xi}_r^2=\frac{(n_1-2S_1)^2\ldots(n_r-2S_r)^2}{T}=\prod_{j=1}^{r}\frac{(n_j-2S_j)^2}{n_j}. \tag{18}$$

Тогда

$$\ln|\tilde{\xi}_r|=\ln\varsigma_1+\ldots+\ln\varsigma_r, \tag{19}$$

где $\varsigma_j=n_j^{-1/2}|n_j-2S_j|,\quad j=1,\ldots,r.$ Центрируя и нормируя обе части формулы (19), получаем, что

$$\frac{\ln|\tilde{\xi}_r|-\mathbf{E}\ln|\tilde{\xi}_r|}{\sqrt{\mathbf{D}\ln|\tilde{\xi}_r|}}=\sum_{j=1}^{r}\frac{\varsigma_j-\mathbf{E}\varsigma_j}{\sqrt{\sum_{k=1}^{r}\mathbf{D}\varsigma_j}}. \tag{20}$$

Так как случайная величина S_j распределена по биномиальному закону с параметрами n_j и $1/2$, то моменты случайных величин ς_j определяются формулами

$$\mathbf{E}\ln^k\varsigma_j=\frac{1}{2^{n_j}}\sum_{l=0}^{n_j}C_{n_j}^{l}\ln^k\left(\frac{|2l-n_j|}{\sqrt{n_j}}\right),\quad k=1,2,\ldots.$$

Если числа n_j нечетны, то величины $\mathbf{E}\ln^k\varsigma_j$ конечны при всех $k,j.$ Кроме того, из сходимости распределения случайной величины $n_j^{-1/2}(2S_j-n_j)$ к стандартному нормальному закону при $n_j\to\infty$ следует, что $\mathbf{E}\ln^k\varsigma_j$ сходится к $\sqrt{\frac{2}{\pi}}\int_0^{+\infty}\ln^k x e^{-x^2/2}dx$, который конечен. Значит, $\mathbf{E}\ln^k\varsigma_j$ равномерно ограничены по $j.$ То же самое относится и к центральным моментам. В частности, $\mathbf{D}\varsigma_j$ сходится к σ^2 равномерно по $j.$ Значит,

$$\mathbf{D}\ln|\tilde{\xi}_r|=r\sigma^2(1+o(1)). \tag{21}$$

К распределению суммы (20) можно применить неравенство Берри-Эссеена для суммы независимых неодинаково распределенных слагаемых, которое будет иметь вид (см. [14])

$$\left|\mathbf{P}\left\{\frac{\ln|\tilde{\xi}_r| - \mathbf{E}\ln|\tilde{\xi}_r|}{\sqrt{\mathbf{D}\ln|\tilde{\xi}_r|}} < x\right\} - \Phi(x)\right| \le \frac{C}{\sqrt{\mathbf{D}\ln|\tilde{\xi}_r|}}. \tag{22}$$

Из формул (17), (21) и (22) получаем, что при $r \to \infty$ имеет место утверждение 3. Утверждение 3 доказано.

О точности аппроксимации в утверждениях 2 и 3

Согласно неравенству (8) для аппроксимации функции распределения случайной величины $\tilde{\eta}_2$ функцией распределения Ψ_r с точностью до 0,05 необходимо, чтобы длины регистров удовлетворяли условию А и имели значения порядка 1000 (Напомним, что $\tilde{\eta}_2 = \tilde{\eta}_3$). Нами был проведено вычисление точного распределения случайной величины $\tilde{\eta}_2$. Результаты опытов показали, что фактически данная точность аппроксимации имеет место, когда длины регистров на порядок меньше.

Рассмотрим случай $r = 2$. Начнем с небольших длин регистров: $n_1 = 19$, $n_2 = 27$. В этом случае оказывается, что $0{,}159 < d(F_{\tilde{\eta}_2}, \Psi_2) < 0{,}160$. Установлено, что наибольшие отклонения $F_{\tilde{\eta}_2}$ и Ψ_2 имеют место для значений случайной величины $\tilde{\eta}_2$, близких к 0. Для остальных значений качество аппроксимации уже является удовлетворительным.

Дальнейшее численное моделирование показало, что точность аппроксимации примерно в 0,05 достигается для длин регистров порядка 100. Здесь мы не будем приводить все значения найденных отклонений, а рассмотрим только небольшой иллюстрирующий пример. При $n_1 = 49, n_2 = 55$ и при $n_1 = 101, n_2 = 113$ величины расстояния в равномерной метрике оказались равными $0{,}0983 < d(F_{\tilde{\eta}_2}, \Psi_2) < 0{,}0984$ и $0{,}0604 < d(F_{\tilde{\eta}_2}, \Psi_2) < 0{,}0604$ соответственно.

Отметим также, что вычисление точного распределения случайной величины $\tilde{\eta}_2$ при длинах регистров порядка нескольких сотен не представляет собой значительную вычислительную трудность в силу формулы (18). Однако для больших длин регистров вычисления становятся затруднительными. В то же время становится возможным использование функции распределения Ψ_2.

Теперь обратимся к случаю, когда число регистров велико. Ограничимся одним примером. Пусть $r = 5$ и $n_1 = 21, n_2 = 23, n_3 = 25, n_4 = 29, n_5 = 31$. В этом случае оказывается, что

$0,0870 < d(F_9, \Phi) < 0,0871,$, и точность аппроксимации является удовлетворительной.

Заключение

В работе изучены свойства распределения числа пар и троек одинаковых знаков в равновероятной мультициклической случайной последовательности, порожденной $r \geq 2$ регистрами взаимно простых длин $n_1, \ldots, n_r$. Получены предельные распределения указанных случайных величин в двух случаях: при фиксированном r и $n_1, \ldots, n_r \to \infty$ и при $r \to \infty$ и нечетных $n_1, \ldots, n_r \to \infty$. В первом случае получена оценка скорости сходимости к предельному распределению в равномерной метрике. Полученные асимптотические распределения существенно отличаются от распределений аналогичных статистик, построенных по последовательности независимых случайных величин.

Литература

1. Pohl P. Description of MCV, a pseudo-random number generator // Scand. Actuar. J. 1976. Vol. 1. P. 1–14.
2. Меженная Н. М., Михайлов В. Г. О распределении числа единиц в выходной последовательности генератора Пола над полем GF(2) // Матем. вопр. криптогр. 2013. Т. 4, № 4. С. 95–107.
3. Меженная Н. М. О распределении числа единиц в двоичной мультициклической последовательности // ПДМ. 2015. Т. 1 (27). С. 69–77.
4. Mezhennaya N. M. Convergence rate estimators for the number of ones in outcome sequence of MCV generator with m-dependent registers items // Sib. Electron. Math. Reports. 2014. Vol. 11. P. 18–25.
5. Зубков А. М., Михайлов В. Г. Предельные распределения случайных величин, связанных с длинными повторениями в последовательности независимых испытаний // Теория вероятн. и примен. 1974. Т. 19, № 1. С. 173–181.
6. Михайлов В. Г. Предельные распределения случайных величин, связанных с многократными длинными повторениями в последовательности независимых испытаний // Теория вероятн. и примен. 1974. Т. 19, № 1. С. 182–187.
7. Зубков А. М., Михайлов В. Г. О повторениях s-цепочек в последовательности независимых величин // Теория вероятн. и примен. 1979. Т. 24, № 2. С. 267–279.
8. Михайлов В. Г., Шойтов А. М. О длинных повторениях цепочек в цепи Маркова // Дискрет. матем. 2014. Т. 26, № 3. С. 79–89.
9. Михайлов В. Г. Оценки точности пуассоновской аппроксимации для распределения числа серий повторений длинных цепочек в цепи Маркова // Дискрет. матем. 2015. Т. 27, № 4. С. 67–78.

10. Михайлов В. Г., Шойтов А. М. Многократные повторения длинных цепочек в конечной цепи Маркова // Матем. вопр. криптогр. 2015. Т. 6, № 3. С. 117–133.

11. Тюрин И. С. Уточнения остаточного члена в теореме Ляпунова // Теория вероятн. и примен. 2011. Т. 56, вып. 3. С. 808–811.

12. Springer M. D., Thompson W. E. The Distribution of Products of Beta, Gamma and Gaussian Random Variables // SIAM J. Appl. Math. 1970. Vol. 18, № 4. P. 721–737.

13. Бейтмен Г., Эрдейи А. Высшие трансцендентные функции. 2-е изд. М.: Наука, 1973. Т. 1. 296 с.

14. Королев В. Ю., Шевцова И. Г. О верхней оценке абсолютной постоянной в неравенстве Берри — Эссеена // Теория вероятн. и ее примен. 2009. Т. 54, № 4. С. 671–695.

ABOUT THE DISTRIBUTION OF THE NUMBERS OF PAIRS AND TRIPLES OF EQUAL SIGNS IN A CYCLE OF BINARY MULTICYCLIC SEQUENCE

Natalya M. Mezhennaya
Cand. Sci. (Phys. and Math.), A/Prof.,
Bauman Moscow State Technical University
5/1 2nd Baumanskaya St., Moscow 105005, Russia
E-mail: natalia.mezhennaya@gmail.com

The article presents the limit distribution of the numbers of pairs and triples of equal signs in a cycle of binary multicyclic random sequence with r registers filled with uniform distributed mutually independent binary random variables in two cases: 1) when the number of registers r is fixed and their lengths tend to infinity; 2) when r also tends to infinity. In the first case using the Berry-Esseen inequality we have derived the uniform estimates of the degree of convergence.

Keywords: multicyclic sequence; normal limit theorem; number of pairs of equal signs; number of triples of equal signs; Berry — Esseen inequality.