Дискретная математика

том 35 выпуск 1 * 2023

УДК 519.237.33 DOI https://doi.org/10.4213/dm1744

Предельное совместное распределение статистик критериев пакета NIST «Monobit Test», «Frequency Test within a Block» и обобщения критерия «Serial Test»

© 2023 г. М. П. Савелов*

Найдено предельное совместное распределение статистик T_1, T_2, T_3 следующих критериев пакета НИСТ: «Monobit Test», «Frequency Test within a Block» и обобщения критерия «Serial Test» в ситуации, когда исследуемая последовательность состоит из независимых случайных величин, имеющих распределение Бернулли с параметром $p = \frac{1}{2}$. Доказано, что T_1 и (T_2, T_3) асимптотически некоррелированы, а T_2 и T_3 асимптотически положительно коррелированы, причем T_1, T_2, T_3 попарно асимптотически зависимы. Доказано, что ковариационная матрица C предельного распределения вектора (T_1, T_2, T_3) удовлетворяет соотношениям $C_{12} = C_{21} = C_{13} = C_{31} = 0$, $C_{23} = C_{32} > 0$. В случае $p \neq \frac{1}{2}$ описано предельное поведение вектора (T_1, T_2, T_3).

Ключевые слова: совместное распределение статистических критериев, пакет критериев NIST, критерий частот, критерий частот в блоках, критерий подпоследовательностей фиксированной длины, асимптотически некоррелированные статистики, асимптотически независимые статистики

1. Введение

Проверка качества генераторов случайных чисел является важной практической задачей. Для ее решения создан ряд инструментов. В частности, для проверки качества генераторов случайных двоичных последовательностей могут быть использованы такие пакеты статистических критериев, как NIST [1], TestU01, Diehard tests и др. Этой теме посвящено множество работ: например, пакет NIST и смежные вопросы обсуждались в [2–19].

Пусть $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ — последовательность независимых случайных величин, имеющих распределение Бернулли $\mathrm{Bern}(p)$, $p \in (0, 1)$. Через H_{p_0} обозначим гипотезу, в соответствии с которой $p = p_0$. Для проверки гипотезы $H_{\frac{1}{2}}$ в пакете NIST предлагается использовать 15 статистических критериев. Мы рассмотрим статистики критерия частот («Monobit Test»), критерия частот в блоках («Frequency Test within a Block») и критерия подпоследовательностей фиксированной длины («Serial Test»).

*Место работы: МГУ им. М. В. Ломоносова, e-mail: savelovmp@gmail.com

Мы будем предполагать, что в критерии «Frequency Test within a Block» используется N блоков, а в критерии «Serial Test» рассматриваются подпоследовательности длин m и $m-1$. Нас будет интересовать случай, когда $N, m \geqslant 2$ фиксированы и n стремится к бесконечности. Отметим, что критерий подпоследовательностей рассматривался, в частности, в [20, 21].

Положим $M = \left[\frac{n}{N}\right]$. Если из исходной последовательности случайных величин $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ отбросить последние $n - NM$ элементов, то оставшиеся элементы разбиваются на N непересекающихся блоков длины M: первый блок состоит из случайных величин $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_M$, второй — из случайных величин $\varepsilon_{M+1}, \varepsilon_{M+2}, \ldots, \varepsilon_{2M}$, и т. д. Положим

$$\pi_j = \frac{1}{M} \sum_{i=(j-1)M+1}^{jM} \varepsilon_i \quad (1 \leqslant j \leqslant N), \qquad S_k = \sum_{i=1}^{k} \varepsilon_i \quad (0 \leqslant k \leqslant n).$$

Индикатор события A будем обозначать символом I_A.

Пусть $(i_1 \ldots i_m) \in \{0,1\}^m$. Количество появлений цепочки $(i_1 \ldots i_m)$ в последовательности $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n-1}, \varepsilon_n, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{m-1}$ обозначим через $\nu_{i_1 \ldots i_m}$. Другими словами,

$$\nu_{i_1 \ldots i_m} = \sum_{i=1}^{n} I_{(\tilde{\varepsilon}_i \ldots \tilde{\varepsilon}_{i+m-1}) = (i_1 \ldots i_m)}, \tag{1}$$

где $\tilde{\varepsilon}_i = \varepsilon_i I_{i \leqslant n} + \varepsilon_{i-n} I_{i > n}$. Величина $\nu_{i_1 \ldots i_{m-1}}$ определяется аналогично.

Пусть G — непустое подмножество $\{0,1\}^{m-1}$. Положим

$$\Psi^2_{m-1}(G) = \frac{2^{m-1}}{n} \sum_{(i_1 \ldots i_{m-1}) \in G} \left(\nu_{i_1 \ldots i_{m-1}} - \frac{n}{2^{m-1}}\right)^2.$$

Пусть $\bar{G} = \{(b_1 \ldots b_m) \in \{0,1\}^m \mid \exists g \in G : (b_1 \ldots b_{m-1}) = g\}$. Другими словами, элементами множества $\bar{G}$ являются элементы множества G, «дополненные» нулем или единицей справа. Положим

$$\Psi^2_m(\bar{G}) = \frac{2^m}{n} \sum_{(i_1 \ldots i_m) \in \bar{G}} \left(\nu_{i_1 \ldots i_m} - \frac{n}{2^m}\right)^2.$$

В частности,

$$\Psi^2_m(\{0,1\}^m) = \frac{2^m}{n} \sum_{i_1 \ldots i_m} \left(\nu_{i_1 \ldots i_m} - \frac{n}{2^m}\right)^2 = \frac{2^m}{n} \sum_{i_1 \ldots i_m} \nu^2_{i_1 \ldots i_m} - n. \tag{2}$$

Положим

$$T_{\text{mon}} = \frac{2S_n - n}{\sqrt{n}}, \quad T_{\text{fr}} = 4M \sum_{j=1}^{N} \left(\pi_j - \frac{1}{2}\right)^2, \quad T_{\text{serial1}}(G) = \Psi^2_m(\bar{G}) - \Psi^2_{m-1}(G),$$
$$T^{\text{NIST}}_{\text{serial1}} = T_{\text{serial1}}(\{0,1\}^{m-1}) = \Psi^2_m(\{0,1\}^m) - \Psi^2_{m-1}(\{0,1\}^{m-1}),$$
$$T^{\text{NIST}}_{\text{serial2}} = \Psi^2_m(\{0,1\}^m) - 2\Psi^2_{m-1}(\{0,1\}^{m-1}) + \Psi^2_{m-2}(\{0,1\}^{m-2}).$$

Статистики $T^{\mathrm{NIST}}_{\mathrm{serial1}}$ и $T^{\mathrm{NIST}}_{\mathrm{serial2}}$ можно определить при всех $m \geqslant 1$ (см. [1]), но для простоты мы будем рассматривать $T^{\mathrm{NIST}}_{\mathrm{serial1}}$ и $T_{\mathrm{serial1}}(G)$ только при $m \geqslant 2$, а $T^{\mathrm{NIST}}_{\mathrm{serial2}}$ — только при $m \geqslant 3$.

Статистики T_{mon}, T_{fr}, $T^{\mathrm{NIST}}_{\mathrm{serial1}}$, $T^{\mathrm{NIST}}_{\mathrm{serial2}}$ используются в критериях согласия «Monobit Test», «Frequency Test within a Block» и «Serial Test» пакета [1] для проверки гипотезы $H_{\frac{1}{2}}$. При этом предполагается, что числа m и N фиксированы и $n \to \infty$. Мы предлагаем новую статистику $T_{\mathrm{serial1}}(G)$, которая является обобщением статистики $T^{\mathrm{NIST}}_{\mathrm{serial1}}$. В настоящей работе основное внимание уделяется тройке статистик T_{mon}, T_{fr} и $T_{\mathrm{serial1}}(G)$.

Предельные распределения отдельных статистик T_{mon}, T_{fr}, $T^{\mathrm{NIST}}_{\mathrm{serial1}}$, $T^{\mathrm{NIST}}_{\mathrm{serial2}}$ известны (см. [1]). Далее, под асимптотической независимостью и асимптотической некоррелированностью статистик будем понимать то же, что понимается под этими терминами в [18,19]. Приведем соответствующие определения для полноты изложения.

Предположим, что многомерные статистики $\vec{T} \in \mathbb{R}^{d_1}$ и $\vec{T}^* \in \mathbb{R}^{d_2}$ построены по выборке $(\varepsilon_1, \ldots, \varepsilon_n)$. Будем говорить, что статистики $\vec{T} \in \mathbb{R}^{d_1}$ и $\vec{T}^* \in \mathbb{R}^{d_2}$ асимптотически независимы, если существуют такие независимые случайные векторы $\vec{\zeta} \in \mathbb{R}^{d_1}$ и $\vec{\zeta}^* \in \mathbb{R}^{d_2}$, что $(\vec{T}, \vec{T}^*) \xrightarrow{d} (\vec{\zeta}, \vec{\zeta}^*)$ при $n \to \infty$. Будем говорить, что статистики $\vec{T} = (T_1, \ldots, T_{d_1}) \in \mathbb{R}^{d_1}$ и $\vec{T}^* = (T^*_1, \ldots, T^*_{d_2}) \in \mathbb{R}^{d_2}$ асимптотически некоррелированы (асимптотически неотрицательно коррелированы), если при всех $1 \leqslant i \leqslant d_1$, $1 \leqslant j \leqslant d_2$ выполнено равенство $\lim_{n\to\infty} \mathrm{cov}(T_i, T^*_j) = 0$ (соответственно, $\underline{\lim}_{n\to\infty} \mathrm{cov}(T_i, T^*_j) \geqslant 0$). Асимптотическая положительная коррелированность определяется аналогично с помощью неравенства $\underline{\lim}_{n\to\infty} \mathrm{cov}(T_i, T^*_j) > 0$.

Обозначим через T_{runs}, T_{template}, T_{cusum}, $T_{\mathrm{longest\ run}}$, T_{matrix} статистики критериев «Runs Test», «Non-overlapping Template Matching Test», «Cumulative Sums Test», «Test for the Longest Run of Ones in a Block» и «Binary Matrix Rank Test» пакета НИСТ. Совместное распределение статистик T_{mon}, T_{fr}, T_{runs} изучалось В. Г. Михайловым в 2019–2020 гг. Совместному распределению статистик T_{mon}, T_{fr}, T_{cusum} посвящена работа [16]. В [15] найдено предельное совместное распределение статистик T_{mon}, T_{fr}, T_{runs}, T_{template}, в [18] — предельное совместное распределение статистик T_{mon}, T_{fr}, $T_{\mathrm{longest\ run}}$, в [19] — предельное совместное распределение статистик T_{mon}, T_{fr}, T_{matrix}.

В данной работе основное внимание уделяется ситуации, когда исследуемая последовательность состоит из независимых случайных величин, имеющих распределение Бернулли с параметром $\frac{1}{2}$. Получены следующие результаты. Для непустых множеств $G_1, \ldots, G_r \subset \{0,1\}^{m-1}$ найдены предельное распределение вектора $(T_{\mathrm{mon}}, T_{\mathrm{fr}}, T_{\mathrm{serial1}}(G_1), \ldots, T_{\mathrm{serial1}}(G_r))$ и ковариационная матрица, соответствующая этому распределению (теорема 2). Как следствие, найдено предельное распределение вектора $(T_{\mathrm{mon}}, T_{\mathrm{fr}}, T^{\mathrm{NIST}}_{\mathrm{serial1}})$ (следствие 3). Получена ковариационная матрица предельного распределения вектора $(T_{\mathrm{mon}}, T_{\mathrm{fr}}, T^{\mathrm{NIST}}_{\mathrm{serial1}}, T^{\mathrm{NIST}}_{\mathrm{serial2}})$ (лемма 2). Доказано, что статистики $T_{\mathrm{mon}}, T_{\mathrm{fr}}, T^{\mathrm{NIST}}_{\mathrm{serial1}}$ попарно асимптотически зависимы (следствие 3). Доказана асимптотическая некоррелированность статистики T_{mon} и вектора $(T^{\mathrm{NIST}}_{\mathrm{serial1}}, T^{\mathrm{NIST}}_{\mathrm{serial2}})$ и асимптотическая некоррелированность T_{fr} и $T^{\mathrm{NIST}}_{\mathrm{serial2}}$ (следствия 3–4). Установлена асимптотическая положительная коррелированность T_{fr} и $T^{\mathrm{NIST}}_{\mathrm{serial1}}$ (следствие 3) и асимптотическая положительная коррелированность $T^{\mathrm{NIST}}_{\mathrm{serial1}}$ и $T^{\mathrm{NIST}}_{\mathrm{serial2}}$ (лемма 2). В лемме 4 описано предельное поведение

вектора $(T_{\text{mon}}, T_{\text{fr}}, T^{\text{NIST}}_{\text{serial1}}, T^{\text{NIST}}_{\text{serial2}})$ в случае, когда при некотором $p \in (0,1)\backslash\{\frac{1}{2}\}$ верна гипотеза H_p.

2. Основные результаты

Занумеруем наборы длины $m-1$ из нулей и единиц в лексикографическом порядке: $B^{(1)} = (00\ldots0), \ldots, B^{(2^{m-1})} = (11\ldots1)$. Если $B = (i_1 \ldots i_{m-1})$ и $j \in \{0,1\}$, то под величиной ν_{Bj} будем понимать $\nu_{i_1\ldots i_{m-1}j}$. Положим

$$Y_n(B) = \frac{2^{\frac{m-1}{2}}}{\sqrt{n}}(\nu_{B1} - \nu_{B0}), \quad B \subset \{0,1\}^{m-1}.$$

Мощность множества G обозначим через $|G|$.

Теорема 1. *Пусть $\varepsilon_1, \varepsilon_2, \ldots$ — последовательность испытаний Бернулли с параметром $p = \frac{1}{2}$. Если числа $m \geqslant 2$ и N фиксированы и $n \to \infty$, то случайные величины $Y_n(B^{(1)}), \ldots, Y_n\big(B^{(2^{m-1})}\big)$ асимптотически независимы и предельным распределением каждого из них является стандартное нормальное распределение $\mathcal{N}(0,1)$. Кроме того,*

$$T_{\text{mon}} = 2^{-\frac{m-1}{2}} \sum_{j=1}^{2^{m-1}} Y_n(B^{(j)}), \tag{3}$$

$$T_{\text{serial1}}(\{B\}) = Y_n^2(B), \quad B \subset \{0,1\}^{m-1}.$$

Если $G \subset \{0,1\}^{m-1}$ и $G \neq \varnothing$, то

$$T_{\text{serial1}}(G) = \sum_{B\in G} T_{\text{serial1}}(\{B\}) = \sum_{B\in G} Y_n^2(B). \tag{4}$$

Как следствие, $\mathcal{L}(T_{\text{mon}}) \to \mathcal{N}(0,1)$ и $\mathcal{L}(T_{\text{serial1}}(G)) \to \chi^2_{|G|}$ при $n \to \infty$. В частности, $\mathcal{L}(T^{\text{NIST}}_{\text{serial1}}) \to \chi^2_{2^{m-1}}$, $n \to \infty$.

Из теоремы 1 следует, что статистика $T_{\text{serial1}}(G)$ является суммой имеющих более простой вид и асимптотически независимых статистик $T_{\text{serial1}}(\{B\})$, $B \in G$, предельным распределением каждой из которых является χ^2_1.

Следствие 1. *Пусть $\varepsilon_1, \varepsilon_2, \ldots$ — последовательность испытаний Бернулли с параметром $p = \frac{1}{2}$. Если числа $m \geqslant 2$ и N, а также непустые непересекающиеся множества $G_i \subset \{0,1\}^{m-1}$, $1 \leqslant i \leqslant r$, фиксированы и $n \to \infty$, то статистики $T_{\text{serial1}}(G_1), \ldots, T_{\text{serial1}}(G_r)$ асимптотически независимы и $\mathcal{L}(T_{\text{serial1}}(G_i)) \to \chi^2_{|G_i|}$, $1 \leqslant i \leqslant r$.*

Теорема 2. *Пусть $\varepsilon_1, \varepsilon_2, \ldots$ — последовательность испытаний Бернулли с параметром $p = \frac{1}{2}$ и $G_1, G_2, \ldots, G_r \subset \{0,1\}^{m-1}$ — непустые множества. Рассмотрим независимые случайные величины $Z_j^{(i)}$, $1 \leqslant i \leqslant N$, $1 \leqslant j \leqslant 2^{m-1}$, имеющие распределение $\mathcal{N}(0,1)$. Положим $Z^{(i)}_{2^{m-1}+1} = 2^{-\frac{m-1}{2}} \sum_{j=1}^{2^{m-1}} Z_j^{(i)}$. Если числа $m \geqslant 2$ и N*

фиксированы, то

$$\left(\sqrt{N}T_{\text{mon}}, T_{\text{fr}}, N\cdot T_{\text{serial1}}(G_1), N\cdot T_{\text{serial1}}(G_2), \ldots, N\cdot T_{\text{serial1}}(G_r)\right) \xrightarrow{d} \left(\sum_{i=1}^{N} Z^{(i)}_{2^{m-1}+1}, \right.$$
$$\left. \sum_{i=1}^{N}\left(Z^{(i)}_{2^{m-1}+1}\right)^2, \sum_{j:B^{(j)}\in G_1}\left(\sum_{i=1}^{N} Z^{(i)}_j\right)^2, \sum_{j:B^{(j)}\in G_2}\left(\sum_{i=1}^{N} Z^{(i)}_j\right)^2, \ldots, \sum_{j:B^{(j)}\in G_r}^{d}\left(\sum_{i=1}^{N} Z^{(i)}_j\right)^2\right) \quad (5)$$

при $n\to\infty$. *Ковариационная матрица предельного распределения вектора* $(T_{\text{mon}}, T_{\text{fr}}, T_{\text{serial1}}(G_i), T_{\text{serial1}}(G_j))$ *имеет вид*

$$F = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2N & |G_i|2^{2-m} & |G_j|2^{2-m} \\ 0 & |G_i|2^{2-m} & 2|G_i| & 2|G_i\cap G_j| \\ 0 & |G_j|2^{2-m} & 2|G_i\cap G_j| & 2|G_j| \end{pmatrix}. \quad (6)$$

При каждом $i=1,2,\ldots,r$ *статистики* T_{mon}, T_{fr}, $T_{\text{serial1}}(G_i)$ *попарно асимптотически зависимы. Кроме того,* $T_{\text{serial1}}(G_i)$ *и* $T_{\text{serial1}}(G_j)$ *асимптотически независимы тогда и только тогда, когда* $G_i\cap G_j=\varnothing$.

Как легко видеть, теорема 2 позволяет получить и предельное распределение вектора $(T_{\text{mon}}, T_{\text{fr}}, T_{\text{serial1}}(G_1), \ldots, T_{\text{serial1}}(G_r))$, и ковариационную матрицу, соответствующую этому распределению.

Лемма 1. *Если выполнены условия теоремы* 2, *то имеет место поэлементная сходимость ковариационных матриц* $\mathbf{D}(T_{\text{mon}}, T_{\text{fr}}, T_{\text{serial1}}(G_i), T_{\text{serial1}}(G_j)) \to F$, $n\to\infty$, *где* F *определена в* (6).

Из теоремы 2 и леммы 1 тривиальным образом получаем следующий результат.

Следствие 2. *Если выполнены условия теоремы* 2, *то при каждом* $1\leqslant i\leqslant r$ *статистики* T_{mon} *и* $(T_{\text{fr}}, T_{\text{serial1}}(G_i))$ *асимптотически некоррелированы, а статистики* T_{fr} *и* $T_{\text{serial1}}(G_i)$ *асимптотически положительно коррелированы.*

Отметим, что теорема 1, в отличие от теоремы 2, ничего не говорит о статистике T_{fr}, однако теорема 1 содержит более простое выражение для предельного совместного распределения статистик T_{mon} и $T_{\text{serial1}}(G)$, так как данное предельное распределение в теореме 1 задается функцией от 2^{m-1} независимых случайных величин со стандартным нормальным распределением, в то время как в теореме 2 для описания этого предельного распределения требуется $N2^{m-1}$ независимых случайных величин со стандартным нормальным распределением.

Подставляя $r=1$ и $G_1=\{0,1\}^{m-1}$ в теорему 2 и следствие 2, получаем следующий результат.

Следствие 3. *Пусть* $\varepsilon_1,\varepsilon_2,\ldots$ — *последовательность испытаний Бернулли с параметром* $p=\frac{1}{2}$. *Пусть случайные величины* $Z^{(i)}_j$, $1\leqslant i\leqslant N$, $1\leqslant j\leqslant 2^{m-1}$, *независимы, имеют распределение* $\mathcal{N}(0,1)$ *и* $Z^{(i)}_{2^{m-1}+1} = 2^{-\frac{m-1}{2}}\sum_{j=1}^{2^{m-1}} Z^{(i)}_j$. *Если* $m\geqslant 2$ *и* N *фиксированы, то*

$$\left(\sqrt{N}T_{\text{mon}}, T_{\text{fr}}, N\cdot T^{\text{NIST}}_{\text{serial1}}\right) \xrightarrow{d} \left(\sum_{i=1}^{N} Z^{(i)}_{2^{m-1}+1}, \sum_{i=1}^{N}\left(Z^{(i)}_{2^{m-1}+1}\right)^2, \sum_{j=1}^{2^{m-1}}\left(\sum_{i=1}^{N} Z^{(i)}_j\right)^2\right)$$

при $n \to \infty$. *Ковариационная матрица предельного распределения вектора* $(T_{\mathrm{mon}}, T_{\mathrm{fr}}, T^{\mathrm{NIST}}_{\mathrm{serial1}})$ *имеет вид*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2N & 2 \\ 0 & 2 & 2^m \end{pmatrix}.$$

Статистики T_{mon}, T_{fr}, $T^{\mathrm{NIST}}_{\mathrm{serial1}}$ *попарно асимптотически зависимы. Статистики* T_{mon} *и* $(T_{\mathrm{fr}}, T^{\mathrm{NIST}}_{\mathrm{serial1}})$ *асимптотически некоррелированы, статистики* T_{fr} *и* $T^{\mathrm{NIST}}_{\mathrm{serial1}}$ *асимптотически положительно коррелированы.*

Отметим, что в силу формул (3), (4), асимптотических свойств $Y_n(B)$ и следствия 3 имеет место следующий результат: если $N = 2^{m-1}$, то предельное распределение вектора $(T_{\mathrm{mon}}, T_{\mathrm{fr}})$ и предельное распределение вектора $(T_{\mathrm{mon}}, T^{\mathrm{NIST}}_{\mathrm{serial1}})$ одинаковы, так как они оба совпадают с распределением вектора $\left(N^{-\frac{1}{2}} \sum_{i=1}^{N} \eta_i, \sum_{i=1}^{N} \eta_i^2\right)$, где случайные величины η_i, $1 \leqslant i \leqslant N$, независимы и имеют стандартное нормальное распределение.

Далее, пусть $m \geqslant 3$ и $G \subset \{0,1\}^{m-1}$ — такое множество, что $G = \{g\delta \mid g \in \underline{G}, \delta \in \{0,1\}\}$ при некотором непустом $\underline{G} \subset \{0,1\}^{m-2}$, где $g\delta$ обозначает $(m-1)$-элементный набор из нулей и единиц, получающийся из g дописыванием справа числа δ. Тогда определены статистики

$$T_{\mathrm{serial1}}(\underline{G}) = \Psi^2_m(G) - \Psi^2_{m-1}(\underline{G}),$$
$$T_{\mathrm{serial2}}(G) = T_{\mathrm{serial1}}(G) - T_{\mathrm{serial1}}(\underline{G}) = \Psi^2_m(\bar{G}) - 2\Psi^2_{m-1}(G) + \Psi^2_{m-2}(\underline{G}).$$

В частности, $T^{\mathrm{NIST}}_{\mathrm{serial2}} = T_{\mathrm{serial2}}(\{0,1\}^{m-1})$.

Следствие 4. *Пусть* $\varepsilon_1, \varepsilon_2, \ldots$ *— последовательность испытаний Бернулли с параметром* $p = \frac{1}{2}$. *Предположим, что* $m \geqslant 3$ *и существует такое* $\underline{G} \subset \{0,1\}^{m-2}$, *что* $G = \{g\delta \mid g \in \underline{G}, \delta \in \{0,1\}\}$. *Если* m, N *и* G *фиксированы, то вектор* $(T_{\mathrm{mon}}, T_{\mathrm{fr}})$ *и статистика* $T_{\mathrm{serial2}}(G)$ *асимптотически некоррелированы при* $n \to \infty$. *В частности, вектор* $(T_{\mathrm{mon}}, T_{\mathrm{fr}})$ *и статистика* $T^{\mathrm{NIST}}_{\mathrm{serial2}}$ *асимптотически некоррелированы при* $n \to \infty$.

Лемма 2. *Пусть* $\varepsilon_1, \varepsilon_2, \ldots$ *— последовательность испытаний Бернулли с параметром* $p = \frac{1}{2}$. *Если* $m \geqslant 3$ *и* N *фиксированы, то матрица*

$$F^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2N & 2 & 0 \\ 0 & 2 & 2^m & 2^{m-1} \\ 0 & 0 & 2^{m-1} & 2^{m-1} \end{pmatrix}$$

является ковариационной матрицей предельного (при $n \to \infty$*) распределения вектора* $(T_{\mathrm{mon}}, T_{\mathrm{fr}}, T^{\mathrm{NIST}}_{\mathrm{serial1}}, T^{\mathrm{NIST}}_{\mathrm{serial2}})$ *и имеет место поэлементная сходимость ковариационных матриц* $\mathbf{D}(T_{\mathrm{mon}}, T_{\mathrm{fr}}, T^{\mathrm{NIST}}_{\mathrm{serial1}}, T^{\mathrm{NIST}}_{\mathrm{serial2}}) \to F^{(2)}$, $n \to \infty$. *Как следствие, статистики* $T^{\mathrm{NIST}}_{\mathrm{serial1}}$ *и* $T^{\mathrm{NIST}}_{\mathrm{serial2}}$ *асимптотически положительно коррелированы.*

Лемма 2 дополняет результат следствия 3 в случае, когда $m \geqslant 3$.

Далее, из [21] легко получить следующую лемму.

Лемма 3. *Пусть* $\varepsilon_1, \varepsilon_2, \ldots$ *— последовательность испытаний Бернулли с параметром* $p = \frac{1}{2}$. *Пусть независимые случайные величины* ξ *и* η *имеют распределение* $\chi^2_{2^{m-2}}$. *Если числа* $m \geqslant 3$ *и* N *фиксированы, то при* $n \to \infty$

$$\left(T^{\mathrm{NIST}}_{\mathrm{serial1}}, T^{\mathrm{NIST}}_{\mathrm{serial2}}\right) \xrightarrow{d} (\xi + \eta, \eta).$$

Замечание 1. Пусть $m \geqslant 3$. В силу леммы 3 статистики $T^{\mathrm{NIST}}_{\mathrm{serial1}}$ и $T^{\mathrm{NIST}}_{\mathrm{serial2}}$ асимптотически зависимы. Можно немного изменить определение $T^{\mathrm{NIST}}_{\mathrm{serial1}}$ так, чтобы добиться асимптотической независимости с $T^{\mathrm{NIST}}_{\mathrm{serial2}}$. А именно, заменив m на $m-1$ в определении статистики $T^{\mathrm{NIST}}_{\mathrm{serial1}}$, получим статистику $T^{*}_{\mathrm{serial1}} = \Psi^2_{m-1}(\{0,1\}^{m-1}) - \Psi^2_{m-2}(\{0,1\}^{m-2})$. Так как $T^{*}_{\mathrm{serial1}} = T^{\mathrm{NIST}}_{\mathrm{serial1}} - T^{\mathrm{NIST}}_{\mathrm{serial2}}$, то из леммы 3 следует, что статистики T^{*}_{serial1} и $T^{\mathrm{NIST}}_{\mathrm{serial2}}$ асимптотически независимы и $\mathcal{L}(T^{*}_{\mathrm{serial1}}) \to \chi^2_{2^{m-2}}$ при $n \to \infty$.
Замечание 2. Пусть выполнены условия следствия 4. Положим

$$\mathring{Y}_n(\underline{B}) = \frac{2^{\frac{m-1}{2}}}{\sqrt{n}}(\nu_{\underline{B}11} - \nu_{\underline{B}01}), \quad \underline{B} \subset \{0,1\}^{m-2}.$$

Можно показать (см. равенства (30), (32) ниже), что

$$T_{\mathrm{serial1}}(G) = \sum_{\underline{B} \in \underline{G}} \big(Y_n^2(\underline{B}1) + Y_n^2(\underline{B}0)\big), \tag{7}$$

$$T_{\mathrm{serial1}}(\underline{G}) = \sum_{\underline{B} \in \underline{G}} T_{\mathrm{serial1}}(\{\underline{B}\}) = \frac{1}{2} \sum_{\underline{B} \in \underline{G}} \big(2\mathring{Y}_n(\underline{B}) - Y_n(\underline{B}1) + Y_n(\underline{B}0)\big)^2. \tag{8}$$

Так как $T_{\mathrm{serial2}}(G) = T_{\mathrm{serial1}}(G) - T_{\mathrm{serial1}}(\underline{G})$, то вектор $(T_{\mathrm{serial1}}(G), T_{\mathrm{serial1}}(\underline{G}), T_{\mathrm{serial2}}(G))$ явным образом выражается через величины $\mathring{Y}_n(\underline{B})$, $Y_n(\underline{B}1)$, $Y_n(\underline{B}0)$. Занумеруем наборы длины $m-2$ из нулей и единиц в лексикографическом порядке: $\underline{B}^{(1)} = (00\ldots0), \ldots, \underline{B}^{(2^{m-2})} = (11\ldots1)$. Предельное распределение вектора

$$\big(\mathring{Y}_n(\underline{B}^{(1)}), Y_n(\underline{B}^{(1)}1), Y_n(\underline{B}^{(1)}0), \ldots, \mathring{Y}_n(\underline{B}^{(2^{m-2})}), Y_n(\underline{B}^{(2^{m-2})}1), Y_n(\underline{B}^{(2^{m-2})}0)\big)$$

может быть найдено из тех же соображений, которые будут использованы при доказательстве соотношения (11) ниже. При этом $T_{\mathrm{mon}} = 2^{-\frac{m-1}{2}} \sum_{j=1}^{2^{m-1}} Y_n(B^{(j)}) = 2^{-\frac{m-1}{2}} \sum_{\underline{B} \in \underline{G}} (Y_n(\underline{B}1) + Y_n(\underline{B}0))$. Далее, предположим, что множества $G_1, G_2, \ldots, G_r \subset \{0,1\}^{m-1}$ не являются пустыми и при каждом $1 \leqslant i \leqslant r$ существует такое $\underline{G}_i \subset \{0,1\}^{m-2}$, что $G_i = \{g\delta \mid g \in \underline{G}_i, \delta \in \{0,1\}\}$. Тогда предельное распределение вектора

$$\big(\sqrt{N}T_{\mathrm{mon}}, T_{\mathrm{fr}}, N \cdot T_{\mathrm{serial1}}(G_1), N \cdot T_{\mathrm{serial2}}(G_1), N \cdot T_{\mathrm{serial1}}(\underline{G}_1), N \cdot T_{\mathrm{serial1}}(G_2),$$
$$N \cdot T_{\mathrm{serial2}}(G_2), N \cdot T_{\mathrm{serial1}}(\underline{G}_2), \ldots, N \cdot T_{\mathrm{serial1}}(G_r), N \cdot T_{\mathrm{serial2}}(G_r), N \cdot T_{\mathrm{serial1}}(\underline{G}_r)\big)$$

может быть найдено из тех же соображений, из которых получено соотношение (5). Отметим также, что данный подход приводит к громоздким промежуточным выкладкам, однако он позволяет, в частности, получить предельное распределение вектора $(T_{\mathrm{mon}}, T_{\mathrm{fr}}, T^{\mathrm{NIST}}_{\mathrm{serial1}}, T^{\mathrm{NIST}}_{\mathrm{serial2}})$ и соответствующую ковариационную матрицу (при $r = 1$ и $G_1 = \{0,1\}^{m-1}$).

Далее, в силу теоремы 2 предельное распределение вектора $(T_{\mathrm{mon}}, T_{\mathrm{fr}}, T_{\mathrm{serial1}}(G))$ известно в случае, когда верна гипотеза $H_{\frac{1}{2}}$, фиксированы числа $m \geqslant 2$, N и множество G и $n \to \infty$. Этот результат дополняет следующая лемма.

Лемма 4. *Пусть $\varepsilon_1, \varepsilon_2, \ldots$ — последовательность испытаний Бернулли с параметром $p \in (0,1)\backslash\{\frac{1}{2}\}$. Если числа $m \geqslant 2$, N и непустое множество $G \subset \{0,1\}^{m-1}$ фиксированы, то*

$$(T_{\mathrm{mon}}, T_{\mathrm{fr}}, T_{\mathrm{serial1}}(G)) \xrightarrow{\text{п.н.}} (\mathrm{sgn}(2p-1) \cdot (+\infty), +\infty, +\infty)$$

при $n \to \infty$. *Если, к тому же,* $m \geqslant 3$ *и существует такое* $\underline{G} \subset \{0,1\}^{m-2}$, *что* $G = \{g\delta \mid g \in \underline{G}, \delta \in \{0,1\}\}$, *то при* $n \to \infty$

$$(T_{\text{mon}}, T_{\text{fr}}, T_{\text{serial1}}(G), T_{\text{serial2}}(G)) \xrightarrow{\text{п.н.}} (\operatorname{sgn}(2p-1)\cdot(+\infty), +\infty, +\infty, +\infty).$$

В частности, $(T_{\text{mon}}, T_{\text{fr}}, T^{\text{NIST}}_{\text{serial1}}, T^{\text{NIST}}_{\text{serial2}}) \xrightarrow{\text{п.н.}} (\operatorname{sgn}(2p-1)\cdot(+\infty), +\infty, +\infty, +\infty)$.

3. Доказательства

3.1. Доказательство теоремы 1. Всюду далее под векторами будем понимать векторы-строки. Если A — матрица, то через $A^\top$ будем обозначать транспонированную матрицу. В [22] (см. формулу (20.59)) сформулирована многомерная центральная предельная теорема для последовательности случайных векторов с ϕ-перемешиванием. Ее частным случаем является следующая лемма.

Лемма 5. *Пусть* $\vec{\zeta}_j$, $j \geqslant 1$, — *стационарная в узком смысле последовательность* m*-зависимых случайных векторов и* $\mathbf{D}\vec{\zeta}_1$ — *ковариационная матрица вектора* $\vec{\zeta}_1$. *Если* $|\zeta_1|$ *меньше некоторой константы с вероятностью* 1, *то матрица* $A = \mathbf{D}\vec{\zeta}_1 + \sum_{k=2}^{m+1}\Big(\operatorname{cov}(\vec{\zeta}_1, \vec{\zeta}_k) + \big(\operatorname{cov}(\vec{\zeta}_1, \vec{\zeta}_k)\big)^\top\Big)$ *симметрична, неотрицательно определена и*

$$\mathcal{L}\left(\frac{\sum_{j=1}^n (\vec{\zeta}_j - \mathbf{E}\vec{\zeta}_j)}{\sqrt{n}}\right) \xrightarrow{d} \mathcal{N}(\vec{0}, A).$$

Следствие 5. *Пусть* $\varepsilon_1, \varepsilon_2, \ldots$ — *последовательность испытаний Бернулли с параметром* $p \in (0,1)$. *Пусть*

$$\vec{\zeta}_j = \big(f_1(\varepsilon_j, \varepsilon_{j+1}, \ldots, \varepsilon_{j+m}), \ldots, f_d(\varepsilon_j, \varepsilon_{j+1}, \ldots, \varepsilon_{j+m})\big),$$

где $f_1, \ldots, f_d$ — *ограниченные борелевские функции. Тогда выполнено утверждение леммы* 5.

При $1 \leqslant j \leqslant n$ положим

$$\vec{\varepsilon}_j^{(s)} = (\varepsilon_j, \varepsilon_{j+1}, \ldots, \varepsilon_{j+s-1}),$$
$$\vec{\zeta}_j = 2^{\frac{m-1}{2}}\big(I_{\vec{\varepsilon}_j^{(m)} = B^{(1)}1} - I_{\vec{\varepsilon}_j^{(m)} = B^{(1)}0}, \ldots, I_{\vec{\varepsilon}_j^{(m)} = B^{(2^{m-1})}1} - I_{\vec{\varepsilon}_j^{(m)} = B^{(2^{m-1})}0}\big).$$

Векторы $\vec{\zeta}_j$ независимы и одинаково распределены. Пусть

$$C^{(k)} = \operatorname{cov}(\vec{\zeta}_1, \vec{\zeta}_k), \quad 1 \leqslant k \leqslant m.$$

В силу следствия 5 имеем:

$$\mathcal{L}\left(\frac{1}{\sqrt{n}}\sum_{j=1}^n (\vec{\zeta}_j - \mathbf{E}\vec{\zeta}_j)\right) \xrightarrow{d} \mathcal{N}\Big(\vec{0}, C^{(1)} + \sum_{k=2}^m \big(C^{(k)} + (C^{(k)})^\top\big)\Big). \tag{9}$$

Если $2 \leqslant k \leqslant m$ и $1 \leqslant i, j \leqslant 2^{m-1}$, то

$$2^{-m+1} C_{i,j}^{(k)} = 2^{-m+1} \big(\operatorname{cov}(\vec{\zeta}_1, \vec{\zeta}_k)\big)_{i,j}$$
$$= \mathbf{E}\Big(\big(I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}1} - I_{\vec{\varepsilon}_1^{(m)} = B_1^{(i)}0}\big)\big(I_{\vec{\varepsilon}_k^{(m)} = B^{(j)}1} - I_{\vec{\varepsilon}_k^{(m)} = B^{(j)}0}\big)\Big)$$
$$= \mathbf{E} I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}1} I_{\vec{\varepsilon}_k^{(m-1)} = B^{(j)}} I_{\varepsilon_{k+m-1}=1} + \mathbf{E} I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}0} I_{\vec{\varepsilon}_k^{(m-1)} = B^{(j)}} I_{\varepsilon_{k+m-1}=0}$$
$$- \mathbf{E} I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}0} I_{\vec{\varepsilon}_k^{(m-1)} = B^{(j)}} I_{\varepsilon_{k+m-1}=1} - \mathbf{E} I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}1} I_{\vec{\varepsilon}_k^{(m-1)} = B^{(j)}} I_{\varepsilon_{k+m-1}=0}$$
$$= \frac{1}{2}\Big(\mathbf{E} I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}1} I_{\vec{\varepsilon}_k^{(m-1)} = B^{(j)}} + \mathbf{E} I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}0} I_{\vec{\varepsilon}_k^{(m-1)} = B^{(j)}} \Big)$$
$$- \frac{1}{2}\Big(\mathbf{E} I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}0} I_{\vec{\varepsilon}_k^{(m-1)} = B^{(j)}} + \mathbf{E} I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}1} I_{\vec{\varepsilon}_k^{(m-1)} = B^{(j)}} \Big) = 0.$$

При $k = 1$, $1 \leqslant i, j \leqslant 2^{m-1}$, выполнены равенства

$$2^{-m+1} C_{i,j}^{(k)} = 2^{-m+1} \big(\operatorname{cov}(\vec{\zeta}_1, \vec{\zeta}_1)\big)_{i,j}$$
$$= \mathbf{E}\Big(\big(I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}1} - I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}0}\big)\big(I_{\vec{\varepsilon}_1^{(m)} = B^{(j)}1} - I_{\vec{\varepsilon}_1^{(m)} = B^{(j)}0}\big)\Big)$$
$$= \mathbf{E}\big(I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}1} I_{\vec{\varepsilon}_1^{(m)} = B^{(j)}1} + I_{\vec{\varepsilon}_1^{(m)} = B^{(i)}0} I_{\vec{\varepsilon}_1^{(m)} = B^{(j)}0} - 0\big) = \frac{1}{2^m} I_{i=j} + \frac{1}{2^m} I_{i=j} = \frac{I_{i=j}}{2^{m-1}}.$$

Следовательно, матрица $C^{(1)} + \sum_{k=2}^{m} (C^{(k)} + (C^{(k)})^\top)$ является единичной, откуда в силу (9) и равенства $\mathbf{E}\vec{\zeta}_1 = \vec{0}$ следует, что

$$\mathcal{L}\Big(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \vec{\zeta}_j \Big) \xrightarrow{d} N(\vec{0}, E), \quad n \to \infty, \tag{10}$$

где E — единичная матрица размера $2^{m-1} \times 2^{m-1}$.

Нам потребуется следующее обозначение. Если последовательность случайных величин η_n сходится к нулю по вероятности, то будем писать, что $\eta_n = o_P(1)$.

В силу (1) выполнено равенство

$$Y_n(B^{(j)}) - \frac{2^{\frac{m-1}{2}}}{\sqrt{n}} \Big(\sum_{i=1}^{n} \Big(I_{\vec{\varepsilon}_i^{(m)} = B^{(j)}1} - I_{\vec{\varepsilon}_i^{(m)} = B^{(j)}0} \Big) \Big) = o_P(1), \quad n \to \infty,$$

при $1 \leqslant j \leqslant 2^{m-1}$, поэтому из (10) следует, что

$$\mathcal{L}\Big(\big(Y_n(B^{(1)}), \ldots, Y_n\big(B^{(2^{m-1})}\big)\big) \Big) \xrightarrow{d} N(\vec{0}, E), \quad n \to \infty. \tag{11}$$

Докажем равенство (3). Заметим, что $\sum_{B \in \{0,1\}^{m-1}} I_{\{\varepsilon_i^{(m-1)} = B\}} = 1$. Поэтому $\sum_{B \in \{0,1\}^{m-1}} I_{\{\varepsilon_i^{(m)} = B1\}} = \sum_{B \in \{0,1\}^{m-1}} I_{\{\varepsilon_i^{(m-1)} = B\}} I_{\{\varepsilon_{i+m-1}=1\}} = I_{\{\varepsilon_{i+m-1}=1\}} = \varepsilon_{i+m-1}$. Аналогично получаем, что $\sum_{B \in \{0,1\}^{m-1}} I_{\{\varepsilon_i^{(m)} = B0\}} = 1 - \varepsilon_{i+m-1}$. Применяя данные соображения к последовательности $\tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \ldots$, где $\tilde{\varepsilon}_i = \varepsilon_i I_{i \leqslant n} + \varepsilon_{i-n} I_{i > n}$, получаем, что $\sum_{B \in \{0,1\}^{m-1}} \nu_{B1} = S_n$ и $\sum_{B \in \{0,1\}^{m-1}} \nu_{B0} = n - S_n$. Значит,

$$2^{-\frac{m-1}{2}} \sum_{j=1}^{2^{m-1}} Y_n(B^{(j)}) = \sum_{j=1}^{2^{m-1}} \frac{1}{\sqrt{n}} \big(\nu_{B^{(j)}1} - \nu_{B^{(j)}0}\big) = T_{\text{mon}}.$$

Формула (3) доказана.

Перейдем к доказательству формулы (4).

Пусть $G = \{g_1, \dots, g_d\}$, где $d = |G|$. Евклидову норму вектора v обозначим через $\|v\|$. Заметим, что

$$\frac{n}{2^m}\Psi^2_m(\bar{G}) = \Big\|\nu_{g_1 1} - \frac{n}{2^m}, \dots, \nu_{g_d 1} - \frac{n}{2^m}\Big\|^2 + \Big\|\nu_{g_1 0} - \frac{n}{2^m}, \dots, \nu_{g_d 0} - \frac{n}{2^m}\Big\|^2. \qquad (12)$$

Из определения величин $\nu_{i_1 \dots i_m}$ следует, что $\nu_{B^{(j)}} = \nu_{B^{(j)}1} + \nu_{B^{(j)}0}$. Поэтому

$$\begin{aligned}\frac{n}{2^{m-1}}\Psi^2_{m-1}(G) &= \Big\|\nu_{g_1} - \frac{n}{2^{m-1}}, \dots, \nu_{g_d} - \frac{n}{2^{m-1}}\Big\|^2 \\ &= \Big\|\Big(\nu_{g_1 1} - \frac{n}{2^m}, \dots, \nu_{g_d 1} - \frac{n}{2^m}\Big) + \Big(\nu_{g_1 0} - \frac{n}{2^m}, \dots, \nu_{g_d 0} - \frac{n}{2^m}\Big)\Big\|^2. \qquad (13)\end{aligned}$$

Так как $2(\|v_1\|^2 + \|v_2\|^2) - \|v_1 + v_2\|^2 = \|v_1 - v_2\|^2$, то из (12) и (13) следует, что

$$\begin{aligned}\frac{n}{2^{m-1}}T_{\text{serial1}}(G) &= \frac{n}{2^{m-1}}\Big(\Psi^2_m(\bar{G}) - \Psi^2_{m-1}(G)\Big) \\ &= \Big\|\Big(\nu_{g_1 1} - \frac{n}{2^m}, \dots, \nu_{g_d 1} - \frac{n}{2^m}\Big) - \Big(\nu_{g_1 0} - \frac{n}{2^m}, \dots, \nu_{g_d 0} - \frac{n}{2^m}\Big)\Big\|^2 \\ &= \frac{n}{2^{m-1}}\big\|\big(Y_n(g_1), \dots, Y_n(g_d)\big)\big\|^2 = \frac{n}{2^{m-1}}\sum_{g\in G} Y^2_n(g).\end{aligned}$$

Таким образом, $T_{\text{serial1}}(G) = \sum_{B\in G} Y^2_n(B)$. Следовательно, $T_{\text{serial1}}(\{B\}) = Y^2_n(B)$ при всех $B \subset \{0,1\}^{m-1}$. Значит, $T_{\text{serial1}}(G) = \sum_{B\in G} Y^2_n(B) = \sum_{B\in G} T_{\text{serial1}}(\{B\})$.

Отметим также, что равенства $T_{\text{serial1}}(G) = \sum_{B\in G} T_{\text{serial1}}(\{B\})$ и $T_{\text{serial1}}(\{B\}) = Y^2_n(B)$ легко получить непосредственной проверкой, если учесть, что $\nu_B = \nu_{B1} + \nu_{B0}$.

Теорема 1 доказана.

3.2. Доказательство теоремы 2. Для $k = 1, \dots, N$ положим

$$\begin{aligned}V^{(k)} = \big(V^{(k)}_1, \dots, V^{(k)}_{2^{m-1}+1}\big) = \frac{2^{\frac{m-1}{2}}}{\sqrt{M}} \sum_{i=M(k-1)+1}^{kM-m+1} \Big(&I_{\vec{\varepsilon}^{(m)}_i = B^{(1)}1} - I_{\vec{\varepsilon}^{(m)}_i = B^{(1)}0}, \dots \\ &\dots, I_{\vec{\varepsilon}^{(m)}_i = B^{(2^{m-1})}1} - I_{\vec{\varepsilon}^{(m)}_i = B^{(2^{m-1})}0}, 2^{-\frac{m-1}{2}}(2\varepsilon_i - 1)\Big).\end{aligned}$$

Заметим, что $V^{(1)} = f(\varepsilon_1, \dots, \varepsilon_M)$, $V^{(2)} = f(\varepsilon_{M+1}, \dots, \varepsilon_{2M}), \dots, V^{(N)} = f(\varepsilon_{M(N-1)+1}, \dots, \varepsilon_{MN})$, поэтому векторы $V^{(1)}, \dots, V^{(N)}$ независимы и одинаково распределены. Будем использовать те же обозначения, которые были введены в ходе доказательства теоремы 1. Заметим, что

$$\big(V^{(1)}_1, \dots, V^{(1)}_{2^{m-1}}\big) = \frac{1}{\sqrt{M}} \sum_{i=1}^{M-m+1} \vec{\zeta}_i = \big(Y_M(B^{(1)}), \dots, Y_M(B^{(2^{m-1})})\big) + o_P(1), \qquad (14)$$

и $V^{(1)}_{2^{m-1}+1} = \frac{2S_M - M}{\sqrt{M}} + o_P(1)$.

Применяя теорему 1 к статистике $T^*_{\text{mon}} = \frac{2S_M - M}{\sqrt{M}}$, построенной по последовательности $\varepsilon_1, \dots, \varepsilon_M$, получаем, что $V^{(1)}_{2^{m-1}+1} = \frac{2S_M - M}{\sqrt{M}} + o_P(1) = T^*_{\text{mon}} + o_P(1) = 2^{-\frac{m-1}{2}} \sum_{j=1}^{2^{m-1}} Y_M(B^{(j)}) + o_P(1)$, поэтому в силу (14)

$$V^{(1)}_{2^{m-1}+1} = 2^{-\frac{m-1}{2}} \sum_{j=1}^{2^{m-1}} V^{(1)}_j + o_P(1). \tag{15}$$

Через C обозначим матрицу размера $(2^{m-1}+1) \times (2^{m-1}+1)$, элементы которой имеют следующий вид:

$$C_{i,j} = \begin{cases} I_{i=j} & \text{при } 1 \leqslant i, j \leqslant 2^{m-1}, \\ 1 & \text{при } i = j = 2^{m-1}+1, \\ 2^{-\frac{m-1}{2}} & \text{иначе.} \end{cases}$$

В силу (11), (14) и (15) получаем, что

$$\mathcal{L}\big(V^{(1)}\big) \to \mathcal{N}(0, C). \tag{16}$$

Пусть случайные векторы $W^{(1)}, \dots, W^{(N)}$ независимы и $W^{(i)} \sim \mathcal{N}(0, C)$. Так как векторы $V^{(1)}, \dots, V^{(N)}$ независимы и одинаково распределены, то из (16) следует, что

$$(V^{(1)}, \dots, V^{(N)}) \xrightarrow{d} (W^{(1)}, \dots, W^{(N)}), \quad n \to \infty, \tag{17}$$

где под вектором $(V^{(1)}, \dots, V^{(N)})$ подразумевается $N(2^{m-1}+1)$-мерный вектор, первые $2^{m-1}+1$ координат которого совпадают с соответствующими координатами вектора $V^{(1)}$, вторые $2^{m-1}+1$ координат — с координатами вектора $V^{(2)}$, и т. д.

Отметим, что из (15) и (17) следует, что $W^{(i)}_{2^{m-1}+1} = 2^{-\frac{m-1}{2}} \sum_{j=1}^{2^{m-1}} W^{(i)}_j$, $1 \leqslant i \leqslant N$. Также данное соотношение можно получить непосредственно из определения векторов $W^{(i)}$, заметив, что $\mathbf{E}\big(W^{(i)}_{2^{m-1}+1} - 2^{-\frac{m-1}{2}} \sum_{j=1}^{2^{m-1}} W^{(i)}_j\big)^2 = 0$.

Далее,

$$\sum_{i=1}^{N} V^{(i)}_{2^{m-1}+1} = \frac{2S_n - n}{\sqrt{M}} + o_P(1) = \sqrt{\frac{n}{M}} T_{\text{mon}} + o_P(1). \tag{18}$$

При всех $1 \leqslant j \leqslant 2^{m-1}$ выполнены равенства

$$\Big(\sum_{i=1}^{N} V^{(i)}_j\Big)^2 = 2^{m-1} \left(\frac{\sum_{i=1}^{n} \big(I_{\vec{\varepsilon}^{\,(m)}_i = B^{(j)}1} - I_{\vec{\varepsilon}^{\,(m)}_i = B^{(j)}0}\big)}{\sqrt{n}} \cdot \frac{\sqrt{n}}{\sqrt{M}} + o_P(1) \right)^2$$
$$= \left(Y_n(B^{(j)}) \sqrt{\frac{n}{M}} + o_P(1) \right)^2 = N\big(Y_n(B^{(j)})\big)^2 + o_P(1). \tag{19}$$

Аналогично получаем, что

$$\sum_{i=1}^{N} \big(V^{(i)}_{2^{m-1}+1}\big)^2 = \left(\frac{\sum_{i=1}^{M} (2\varepsilon_i - 1)}{\sqrt{M}} + o_P(1) \right)^2 + \left(\frac{\sum_{i=M+1}^{2M} (2\varepsilon_i - 1)}{\sqrt{M}} + o_P(1) \right)^2 + \cdots$$
$$= T_{\text{fr}} + o_P(1). \tag{20}$$

Из (17)–(20) следует, что

$$\begin{aligned}&\big(\sqrt{N}T_{\text{mon}}, T_{\text{fr}}, NY_n^2\big(B^{(1)}\big), \dots, NY_n^2\big(B^{(2^{m-1})}\big)\big)\\ &\qquad\xrightarrow{d}\left(\sum_{i=1}^N W^{(i)}_{2^{m-1}+1}, \sum_{i=1}^N \big(W^{(i)}_{2^{m-1}+1}\big)^2, \left(\sum_{i=1}^N W_1^{(i)}\right)^2, \dots, \left(\sum_{i=1}^N W^{(i)}_{2^{m-1}}\right)^2\right).\end{aligned}$$

Значит,

$$\begin{aligned}&\big(\sqrt{N}T_{\text{mon}}, T_{\text{fr}}, N\cdot T_{\text{serial1}}(G_1), N\cdot T_{\text{serial1}}(G_2), \dots, N\cdot T_{\text{serial1}}(G_r)\big)\\ &\xrightarrow{d}\left(\sum_{i=1}^N W^{(i)}_{2^{m-1}+1}, \sum_{i=1}^N \big(W^{(i)}_{2^{m-1}+1}\big)^2, \sum_{j:B^{(j)}\in G_1}\left(\sum_{i=1}^N W_j^{(i)}\right)^2, \sum_{j:B^{(j)}\in G_2}\left(\sum_{i=1}^N W_j^{(i)}\right)^2, \dots\right.\\ &\qquad\left.\dots, \sum_{j:B^{(j)}\in G_r}^d\left(\sum_{i=1}^N W_j^{(i)}\right)^2\right). \quad (21)\end{aligned}$$

Так как случайные векторы $W^{(1)}, \dots, W^{(N)}$ независимы и $W^{(i)} \sim \mathcal{N}(0, C)$, то случайные величины $W_j^{(i)}$, $1 \leqslant i \leqslant N$, $1 \leqslant j \leqslant 2^{m-1}$, независимы и имеют стандартное нормальное распределение. Учитывая, что $W^{(i)}_{2^{m-1}+1} = 2^{-\frac{m-1}{2}} \sum_{j=1}^{2^{m-1}} W_j^{(i)}$, из (21) получаем (5).

Найдем явный вид ковариационной матрицы $F = (F_{rs})$ предельного распределения вектора $(T_{\text{mon}}, T_{\text{fr}}, T_{\text{serial1}}(G_i), T_{\text{serial1}}(G_j))$. Предельное распределение каждой из статистик T_{mon}, T_{fr}, $T_{\text{serial1}}(G_i)$, $T_{\text{serial1}}(G_j)$ известно (см. теорему 1 и [1]), что позволяет получить диагональные элементы матрицы F. Далее, пусть координаты вектора $\big(v^{(1)}, v^{(2)}, \dots, v^{(2^{m-1})}\big)$ независимы и имеют стандартное нормальное распределение. В силу теоремы 1 имеем:

$$\begin{aligned}\big(Y_n(B^{(1)}), \dots, Y_n\big(B^{(2^{m-1})}\big)\big) &\xrightarrow{d} \big(v^{(1)}, v^{(2)}, \dots, v^{(2^{m-1})}\big),\\ (T_{\text{serial1}}(G_i), T_{\text{serial1}}(G_j)) &\xrightarrow{d}\left(\sum_{k:B^{(k)}\in G_i}(v^{(k)})^2, \sum_{t:B^{(t)}\in G_j}(v^{(t)})^2\right),\end{aligned}$$

поэтому $F_{34} = F_{43} = \operatorname{cov}\big(\sum_{k:B^{(k)}\in G_i}(v^{(k)})^2, \sum_{t:B^{(t)}\in G_j}(v^{(t)})^2\big) = 2|G_i \cap G_j|$.

Пусть F^* — ковариационная матрица предельного распределения вектора $\big(\sqrt{N}T_{\text{mon}}, T_{\text{fr}}, N\cdot T_{\text{serial1}}(G_i), N\cdot T_{\text{serial1}}(G_j)\big)$. В силу (5) она совпадает с ковариационной матрицей вектора

$$\left(\sum_{k=1}^N Z^{(k)}_{2^{m-1}+1}, \sum_{k=1}^N \big(Z^{(k)}_{2^{m-1}+1}\big)^2, \sum_{s:B^{(s)}\in G_i}\left(\sum_{t=1}^N Z_s^{(t)}\right)^2, \sum_{s:B^{(s)}\in G_j}\left(\sum_{t=1}^N Z_s^{(t)}\right)^2\right).$$

Лемма 6. *Пусть (ξ_1, ξ_2) — двумерный случайный вектор, имеющий нормальное распределение с нулевым средним. Тогда* $\operatorname{cov}(\xi_1, \xi_2^2) = 0$ *и* $\operatorname{cov}(\xi_1^2, \xi_2^2) = 2\operatorname{cov}^2(\xi_1, \xi_2)$.

Применяя лемму 6 к вектору $\left(Z^{(k)}_{2^{m-1}+1}, \sum_{t=1}^{N} Z^{(t)}_s\right)$, получаем, что

$$F^*_{13} = \operatorname{cov}\left(\sum_{k=1}^{N} Z^{(k)}_{2^{m-1}+1}, \sum_{s:B^{(s)}\in G_i}\left(\sum_{t=1}^{N} Z^{(t)}_s\right)^2\right) = \sum_{k=1}^{N}\sum_{s:B^{(s)}\in G_i} \operatorname{cov}\left(Z^{(k)}_{2^{m-1}+1}, \left(\sum_{t=1}^{N} Z^{(t)}_s\right)^2\right) = 0.$$

Аналогичным образом находим, что $F^*_{12} = 0$. Далее,

$$\begin{aligned} F^*_{23} &= \operatorname{cov}\left(\sum_{k=1}^{N}\left(Z^{(k)}_{2^{m-1}+1}\right)^2, \sum_{s:B^{(s)}\in G_i}\left(\sum_{t=1}^{N} Z^{(t)}_s\right)^2\right) \\ &= \sum_{k=1}^{N}\sum_{s:B^{(s)}\in G_i} \operatorname{cov}\left(\left(Z^{(k)}_{2^{m-1}+1}\right)^2, \left(\sum_{t=1}^{N} Z^{(t)}_s\right)^2\right) \\ = \sum_{k=1}^{N}\sum_{s:B^{(s)}\in G_i} 2\operatorname{cov}^2\left(Z^{(k)}_{2^{m-1}+1}, \sum_{t=1}^{N} Z^{(t)}_s\right) &= 2\sum_{k=1}^{N}\sum_{s:B^{(s)}\in G_i} \operatorname{cov}^2\left(Z^{(k)}_{2^{m-1}+1}, Z^{(k)}_s\right) \\ = 2N \sum_{s:B^{(s)}\in G_i} \operatorname{cov}^2\left(Z^{(k)}_{2^{m-1}+1}, Z^{(k)}_s\right) &= 2N \sum_{s:B^{(s)}\in G_i} \left(2^{-\frac{m-1}{2}}\right)^2 = 2^{2-m}N|G_i|. \end{aligned}$$

Так как $F^*_{12} = F^*_{13} = 0$ и $F^*_{23} = 2^{2-m}N|G_i|$, то $F_{12} = F_{13} = 0$ и $F_{23} = 2^{2-m}|G_i|$. Значит,

$$\begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2N & 2^{2-m}|G_i| \\ 0 & 2^{2-m}|G_i| & 2|G_i| \end{pmatrix}.$$

Аналогично получаем, что $F_{14} = F_{41} = 0$, $F_{24} = F_{42} = 2^{2-m}|G_j|$. Следовательно, ковариационная матрица предельного распределения вектора $(T_{\text{mon}}, T_{\text{fr}}, T_{\text{serial1}}(G_i), T_{\text{serial1}}(G_j))$ имеет вид (6).

Покажем, что при каждом $1 \leqslant i \leqslant r$ статистики T_{mon}, T_{fr}, $T_{\text{serial1}}(G_i)$ попарно асимптотически зависимы.

Асимптотическая зависимость T_{mon} и T_{fr} доказана в [18]. Если бы T_{fr} была асимптотически независима с $T_{\text{serial1}}(G_i)$, то выполнялось бы равенство $F_{23} = 0$, однако $F_{23} > 0$.

Предположим, что T_{mon} и $T_{\text{serial1}}(G_i)$ асимптотически независимы. В силу доказанного выше

$$\left(\sqrt{N}T_{\text{mon}}, N \cdot T_{\text{serial1}}(G_i)\right) \xrightarrow{d} \left(\sum_{s=1}^{N} Z^{(s)}_{2^{m-1}+1}, \sum_{j:B^{(j)}\in G_i}\left(\sum_{t=1}^{N} Z^{(t)}_j\right)^2\right). \tag{22}$$

Из (22) и асимптотической независимости T_{mon} и $T_{\text{serial1}}(G_i)$ следует, что случайные величины $\sum_{s=1}^{N} Z^{(s)}_{2^{m-1}+1}$ и $\sum_{j:B^{(j)}\in G_i}\left(\sum_{t=1}^{N} Z^{(t)}_j\right)^2$ независимы. Значит,

$$0 = \operatorname{cov}\left(\left(\sum_{s=1}^{N} Z^{(s)}_{2^{m-1}+1}\right)^2, \sum_{j:B^{(j)}\in G_i}\left(\sum_{t=1}^{N} Z^{(t)}_j\right)^2\right) = \sum_{j:B^{(j)}\in G_i} A_j,$$

где $A_j = \operatorname{cov}\Big(\big(\sum_{s=1}^N Z^{(s)}_{2^{m-1}+1}\big)^2, \big(\sum_{t=1}^N Z^{(t)}_j\big)^2\Big)$. Учитывая лемму 6, получаем, что

$$A_j = 2\operatorname{cov}^2\Big(\sum_{s=1}^N Z^{(s)}_{2^{m-1}+1},\ \sum_{t=1}^N Z^{(t)}_j\Big) = 2\Big(\sum_{s,t=1}^N \operatorname{cov}\Big(Z^{(s)}_{2^{m-1}+1},\ Z^{(t)}_j\Big)\Big)^2$$
$$= 2\Big(\sum_{s=1}^N \operatorname{cov}\Big(Z^{(s)}_{2^{m-1}+1},\ Z^{(s)}_j\Big)\Big)^2 = 2\big(N2^{-\frac{m-1}{2}}\big)^2 > 0.$$

Значит,

$$0 = \operatorname{cov}\left(\left(\sum_{s=1}^N Y^{(s)}_{K+2}\right)^2,\ \sum_{j=1}^{K+1}\left(\sum_{i=1}^N Y^{(i)}_j\right)^2\right) = \sum_{j:B^{(j)}\in G_i} A_j > 0.$$

Полученное противоречие доказывает, что T_{mon} и $T_{\text{serial1}}(G_i)$ асимптотически зависимы.

Осталось доказать, что $T_{\text{serial1}}(G_i)$ и $T_{\text{serial1}}(G_j)$ асимптотически независимы тогда и только тогда, когда $G_i \cap G_j = \varnothing$.

Если $G_i \cap G_j = \varnothing$, то асимптотическая независимость $T_{\text{serial1}}(G_i)$ и $T_{\text{serial1}}(G_j)$ следует из формулы (5).

Если $T_{\text{serial1}}(G_i)$ и $T_{\text{serial1}}(G_j)$ асимптотически независимы, то в силу (6) выполнено равенство $|G_i \cap G_j| = 0$, т. е. $G_i \cap G_j = \varnothing$.

Тем самым теорема 2 доказана.

3.3. Доказательство леммы 1. Нам потребуется несколько вспомогательных утверждений.

В [18] установлен следующий результат.

Лемма 7. *Пусть последовательность двумерных случайных векторов* (α_n, β_n) *сходится по распределению к случайному вектору* (α, β) *при* $n \to \infty$. *Если при некотором* $n_0 \geqslant 1$ *выполнены соотношения* $\sup_{n\geqslant n_0} \mathbf{E}\alpha_n^4 < \infty$ *и* $\sup_{n\geqslant n_0} \mathbf{E}\beta_n^4 < \infty$, *то существуют ковариационные матрицы* $\mathbf{D}(\alpha, \beta)$ *и* $\mathbf{D}(\alpha_n, \beta_n)$, $n \geqslant n_0$, *и имеет место поэлементная сходимость этих матриц*: $\mathbf{D}(\alpha_n, \beta_n) \to \mathbf{D}(\alpha, \beta)$ *при* $n \to \infty$.

Положим $n_0 = \max(N, m)$. Заметим, что каждая из статистик $T_{\text{serial1}}(G)$, $G \subset \{0,1\}^{m-1}$, а также статистики T_{mon} и T_{fr} определены при $n \geqslant n_0$.

Рассуждая так же, как в [18], получаем, что если $\varepsilon_1, \varepsilon_2, \ldots$ — последовательность испытаний Бернулли с параметром $p = \frac{1}{2}$, то

$$\sup_{n\geqslant n_0} \mathbf{E}(T_{\text{mon}})^4 < \infty, \quad \sup_{n\geqslant n_0} \mathbf{E}(T_{\text{fr}})^4 < \infty. \tag{23}$$

Нам осталось доказать, что

$$\sup_{n\geqslant n_0} \mathbf{E}(T_{\text{serial1}}(G))^4 < \infty \tag{24}$$

при всех непустых множествах $G \subset \{0,1\}^{m-1}$. В самом деле, если это будет доказано, то в силу леммы 7 и соотношений (23), (24) мы получим, что выполняется утверждение леммы 1.

Для доказательства (24) нам потребуется следующая лемма.

Лемма 8. *Пусть* ε_i, $1 \leqslant i \leqslant n$, — *последовательность испытаний Бернулли с параметром* $p = \frac{1}{2}$ *и* $f\colon \mathbb{R}^m \to [-L, L]$ — *ограниченная функция. Положим* $\tilde{\varepsilon}_i = \varepsilon_i I_{i\leqslant n} + \varepsilon_{i-n} I_{i>n}$. *Если* $\mathbf{E} f(\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_m) = 0$, *то* $\mathbf{E}\big(\sum_{i=1}^n f(\tilde{\varepsilon}_i, \dots, \tilde{\varepsilon}_{i+m-1})\big)^8 = O(n^4)$, $n \to \infty$.

Доказательство. Положим $\zeta_i = f(\tilde{\varepsilon}_i \dots \tilde{\varepsilon}_{i+m-1})$, $1 \leqslant i \leqslant n$. Из условия леммы следует, что $\mathbf{E}\zeta_i = 0$. Выполнены равенства

$$\mathbf{E}\Big(\sum_{i=1}^n \zeta_i\Big)^8 = \mathbf{E}\Big(\sum_{i_1=1}^n \zeta_{i_1}\Big) \cdots \Big(\sum_{i_8=1}^n \zeta_{i_8}\Big) = \sum_{1\leqslant i_1 \dots i_8 \leqslant n} a_{i_1 \dots i_8}, \tag{25}$$

где $a_{i_1 \dots i_8} = \mathbf{E}\zeta_{i_1} \dots \zeta_{i_8}$. Будем называть индексы i_k и i_s близкими, если множество $\{\tilde{\varepsilon}_{i_k}, \dots, \tilde{\varepsilon}_{i_k+m-1}\}$ пересекается с множеством $\{\tilde{\varepsilon}_{i_s}, \dots, \tilde{\varepsilon}_{i_s+m-1}\}$. Заметим, что если среди индексов $i_1, \dots, i_8$ есть такой индекс i_k, который не близок ни к одному из остальных индексов, то $a_{i_1 \dots i_8} = 0$. Другими словами, если $a_{i_1 \dots i_8} \neq 0$, то у каждого из восьми индексов $i_1, \dots, i_8$ есть близкий к нему. Значит, в сумме $\sum_{1\leqslant i_1 \dots i_8 \leqslant n} a_{i_1 \dots i_8}$ не более $O(n^4)$ ненулевых слагаемых, каждое из которых ограничено величиной L^8. Учитывая (25), получаем, что $\mathbf{E}\big(\sum_{i=1}^n \zeta_i\big)^8 = O(n^4)$, $n \to \infty$. □

Вернемся к доказательству леммы 1. Мы предполагаем, что $\varepsilon_1, \varepsilon_2, \dots$ — последовательность испытаний Бернулли с параметром $p = \frac{1}{2}$. Напомним, что $\nu_{i_1 \dots i_m} = \sum_{i=1}^n I_{(\tilde{\varepsilon}_i \dots \tilde{\varepsilon}_{i+m-1}) = (i_1 \dots i_m)}$. Из леммы 8 следует, что $\sup_{n\geqslant n_0} \mathbf{E} n^{-4}\big(\nu_{i_1 \dots i_m} - \frac{n}{2^m}\big)^8 < \infty$ при всех $i_1, \dots, i_m \in \{0, 1\}$. Применяя неравенство Минковского $\big(\mathbf{E}\big(\sum_{i=1}^K \alpha_i\big)^4\big)^{\frac{1}{4}} \leqslant \sum_{i=1}^K (\mathbf{E}\alpha_i^4)^{\frac{1}{4}}$ к случайным величинам $\frac{1}{n}\big(\nu_{i_1 \dots i_m} - \frac{n}{2^m}\big)^2$ и учитывая, что $\Psi_m(\bar{G}) = \frac{2^m}{n} \sum_{(i_1 \dots i_m) \in \bar{G}} \big(\nu_{i_1 \dots i_m} - \frac{n}{2^m}\big)^2$, получаем: $\sup_{n\geqslant n_0} \mathbf{E}\Psi_m^4(\bar{G}) < \infty$. Рассуждая аналогично, получаем, что $\sup_{n\geqslant n_0} \mathbf{E}\Psi_{m-1}^4(G) < \infty$. В силу неравенства Минковского отсюда следует (24). Лемма 1 доказана.

Заметим, что из аналогичных соображений следует, что

$$\sup_{n\geqslant n_0} \mathbf{E} T_{\text{serial2}}^4(G) < \infty \tag{26}$$

для всех множеств G, для которых определена статистика $T_{\text{serial2}}(G)$.

3.4. Доказательство следствия 4. В силу замечания 2 существует предельное распределение вектора $(T_{\text{mon}}, T_{\text{fr}}, T_{\text{serial2}}(G), T_{\text{serial1}}(G), T_{\text{serial1}}(\underline{G}))$. Обозначим его ковариационную матрицу через $\tilde{F}$.

В силу теоремы 2 выполнено равенство $\tilde{F}_{24} = |G| 2^{2-m}$. Применяя теорему 2 к множеству $\underline{G}$ вместо G и меняя m на $m-1$, получаем, что $\tilde{F}_{25} = |\underline{G}| 2^{2-(m-1)}$. Так как $|G| = 2|\underline{G}|$, то $\tilde{F}_{25} = |G| 2^{2-m}$.

В силу того, что $T_{\text{serial2}}(G) = T_{\text{serial1}}(G) - T_{\text{serial1}}(\underline{G})$, выполнено равенство $\tilde{F}_{23} = \tilde{F}_{24} - \tilde{F}_{25} = 0$. Аналогично получаем, что $\tilde{F}_{13} = 0$.

В силу соотношений (23), (26) и леммы 7 получаем, что асимптотическая некоррелированность $(T_{\text{mon}}, T_{\text{fr}})$ и $T_{\text{serial2}}(G)$ следует из равенств $\tilde{F}_{13} = \tilde{F}_{23} = 0$.

Тем самым следствие 4 доказано.

3.5. Доказательство леммы 2. Положим $G = \{0,1\}^{m-1}$ и рассмотрим матрицу $\tilde{F}$, определенную в доказательстве следствия 4. Матрица $\tilde{F}$ является ковариационной матрицей предельного распределения вектора $(T_{\mathrm{mon}},\ T_{\mathrm{fr}},\ T^{\mathrm{NIST}}_{\mathrm{serial2}},\ T^{\mathrm{NIST}}_{\mathrm{serial1}},\ T_{\mathrm{serial1}}(\{0,1\}^{m-2}))$, и в ходе доказательства следствия 4 установлено, что $\tilde{F}_{13} = \tilde{F}_{23} = 0$, $\tilde{F}_{24} = |G|2^{2-m} = 2$. В силу теоремы 2 выполнены равенства $\tilde{F}_{12} = \tilde{F}_{14} = 0$. Предельные распределения каждой из статистик T_{mon}, T_{fr}, $T^{\mathrm{NIST}}_{\mathrm{serial2}}$ и $T^{\mathrm{NIST}}_{\mathrm{serial1}}$ известны — см. [1]. Следовательно, $\tilde{F}_{11} = 1$, $\tilde{F}_{22} = 2N$, $\tilde{F}_{33} = 2^{m-1}$, $\tilde{F}_{44} = 2^m$.

Далее, в силу (2) мы можем считать, что величина $\Psi^2_k = \Psi^2_k(\{0,1\}^k)$ определена при $k \geqslant 1$. Положим $\Psi^2_0 = \Psi^2_{-1} = 0$, $\nabla\Psi^2_k = \Psi^2_k - \Psi^2_{k-1}$ $(k \geqslant 0)$, $\nabla^2\Psi^2_k = \nabla\Psi^2_k - \nabla\Psi^2_{k-1} = \Psi^2_k - 2\Psi^2_{k-1} + \Psi^2_{k-2}$ $(k \geqslant 1)$. Несложно видеть, что $\nabla\Psi^2_k = \sum_{j=1}^m \nabla^2\Psi^2_j$ $(k \geqslant 1)$.

Мы предполагаем, что $m \geqslant 3$. Выполнены равенства $T^{\mathrm{NIST}}_{\mathrm{serial1}} = \nabla\Psi^2_m$, $T^{\mathrm{NIST}}_{\mathrm{serial2}} = \nabla^2\Psi^2_m$, и, следовательно,

$$\bigl(T^{\mathrm{NIST}}_{\mathrm{serial1}}, T^{\mathrm{NIST}}_{\mathrm{serial2}}\bigr) = \biggl(\sum_{j=1}^m \nabla^2\Psi^2_j, \nabla^2\Psi^2_m\biggr). \tag{27}$$

Как отмечено в [21, с. 104], случайные величины $\nabla^2\Psi^2_1, \nabla^2\Psi^2_2, \ldots$ асимптотически независимы, и если $k \geqslant 2$, то

$$\mathcal{L}\bigl(\nabla^2\Psi^2_k\bigr) \to \chi^2_{2^{k-2}} \tag{28}$$

при $n \to \infty$ (ср. также [24, с. 144]). Отсюда и из (27) следует, что $\tilde{F}_{34} = 2^{m-1}$.

Мы нашли $\tilde{F}_{ij}$ при $1 \leqslant i, j \leqslant 4$. Выражая ковариационную матрицу предельного (при $n \to \infty$) распределения вектора $(T_{\mathrm{mon}}, T_{\mathrm{fr}}, T^{\mathrm{NIST}}_{\mathrm{serial1}}, T^{\mathrm{NIST}}_{\mathrm{serial2}})$ через $\tilde{F}_{ij}$, $1 \leqslant i, j \leqslant 4$, получаем, что она совпадает с $F^{(2)}$. Осталось заметить, что в силу соотношений (23), (24), (26) и леммы 7 имеет место поэлементная сходимость ковариационных матриц $\mathbf{D}(T_{\mathrm{mon}}, T_{\mathrm{fr}}, T^{\mathrm{NIST}}_{\mathrm{serial1}}, T^{\mathrm{NIST}}_{\mathrm{serial2}}) \to F^{(2)}$, $n \to \infty$.

3.6. Доказательство леммы 3. Пусть величины Ψ^2_k, $k \geqslant -1$, те же, что и в доказательстве леммы 2. Заметим, что если $k \geqslant 1$, то

$$\mathcal{L}\bigl(\nabla\Psi^2_k\bigr) \to \chi^2_{2^{k-1}} \tag{29}$$

при $n \to \infty$ согласно [1]. Из равенства (27) следует, что

$$\bigl(T^{\mathrm{NIST}}_{\mathrm{serial1}} - T^{\mathrm{NIST}}_{\mathrm{serial2}}, T^{\mathrm{NIST}}_{\mathrm{serial2}}\bigr) = \biggl(\sum_{j=1}^{m-1} \nabla^2\Psi^2_j, \nabla^2\Psi^2_m\biggr) = \bigl(\nabla\Psi^2_{m-1}, \nabla^2\Psi^2_m\bigr),$$

откуда в силу (28), (29) и асимптотической независимости случайных величин $\nabla^2\Psi^2_1, \nabla^2\Psi^2_2, \ldots$ получаем, что $\bigl(T^{\mathrm{NIST}}_{\mathrm{serial1}} - T^{\mathrm{NIST}}_{\mathrm{serial2}}, T^{\mathrm{NIST}}_{\mathrm{serial2}}\bigr) \xrightarrow{d} (\xi, \eta)$. Следовательно, $\bigl(T^{\mathrm{NIST}}_{\mathrm{serial1}}, T^{\mathrm{NIST}}_{\mathrm{serial2}}\bigr) \xrightarrow{d} (\xi + \eta, \eta)$.

3.7. Доказательство равенств (7)**,** (8) **из замечания 2.** Пусть выполнены условия следствия 4. В силу равенства (4) имеем:

$$T_{\mathrm{serial1}}(G) = \sum_{B \in G} Y^2_n(B) = \sum_{\underline{B} \in \underline{G}} \bigl(Y^2_n(\underline{B}1) + Y^2_n(\underline{B}0)\bigr). \tag{30}$$

Подставляя $m-1$ и $\underline{G}$ вместо m и G в (4) и учитывая равенство $\nu_{\underline{B}\theta} = \nu_{\underline{B}\theta 1} + \nu_{\underline{B}\theta 0}$, $\theta \in \{0,1\}$, получаем, что

$$T_{\mathrm{serial1}}(\underline{G}) = \sum_{\underline{B}\in\underline{G}} \left(\frac{2^{\frac{m-2}{2}}}{\sqrt{n}} \big(\nu_{\underline{B}1} - \nu_{\underline{B}0}\big) \right)^2 = \sum_{\underline{B}\in\underline{G}} \left(\frac{2^{\frac{m-2}{2}}}{\sqrt{n}} \big(\nu_{\underline{B}11} + \nu_{\underline{B}10} - \nu_{\underline{B}01} - \nu_{\underline{B}00}\big) \right)^2$$

$$= \frac{1}{2} \sum_{\underline{B}\in\underline{G}} \left(\frac{2^{\frac{m-1}{2}}}{\sqrt{n}} \big(2(\nu_{\underline{B}11} - \nu_{\underline{B}01}) - (\nu_{\underline{B}11} - \nu_{\underline{B}10}) + (\nu_{\underline{B}01} - \nu_{\underline{B}00})\big) \right)^2$$

$$= \frac{1}{2} \sum_{\underline{B}\in\underline{G}} \big(2\mathring{Y}_n(\underline{B}) - Y_n(\underline{B}1) + Y_n(\underline{B}0)\big)^2. \quad (31)$$

Отсюда следует, что $T_{\mathrm{serial1}}(\{\underline{B}\}) = \frac{1}{2}\big(2\mathring{Y}_n(\underline{B}) - Y_n(\underline{B}1) + Y_n(\underline{B}0)\big)^2$ при всех $\underline{B} \subset \{0,1\}^{m-2}$. Значит,

$$T_{\mathrm{serial1}}(\underline{G}) = \frac{1}{2} \sum_{\underline{B}\in\underline{G}} \big(2\mathring{Y}_n(\underline{B}) - Y_n(\underline{B}1) + Y_n(\underline{B}0)\big)^2 = \sum_{\underline{B}\in\underline{G}} T_{\mathrm{serial1}}(\{\underline{B}\}). \quad (32)$$

3.8. Доказательство леммы 4. Соотношение

$$(T_{\mathrm{mon}}, T_{\mathrm{fr}}) \xrightarrow{\text{п.н.}} (\mathrm{sgn}(2p-1)\cdot(+\infty), +\infty)$$

очевидно (ср. [15, лемма 4]).

Пусть $B \in G$. Положим $p_B = \mathbf{P}((\varepsilon_1, \dots, \varepsilon_{m-1}) = B)$. В силу усиленного закона больших чисел для неотрицательных ограниченных m-зависимых случайных величин $I_{\varepsilon_i^{(m)}=B1}$ (см., например, теорему 1 в [23]) имеем: $\frac{1}{n}\sum_{i=1}^n I_{\varepsilon_i^{(m)}=B1} \xrightarrow{\text{п.н.}} p_B p$. Значит, $\frac{\nu_{B1}}{n} \xrightarrow{\text{п.н.}} p_B p$. Аналогичным образом получаем, что $\frac{\nu_{B0}}{n} \xrightarrow{\text{п.н.}} p_B(1-p)$. Следовательно,

$$Y_n^2(B) = \frac{2^{m-1}}{n}\big(\nu_{B1} - \nu_{B0}\big)^2 \xrightarrow{\text{п.н.}} +\infty.$$

В силу (4) выполнено равенство $T_{\mathrm{serial1}}(G) = \sum_{B\in G} Y_n^2(B)$. Значит, $T_{\mathrm{serial1}}(G) \xrightarrow{\text{п.н.}} +\infty$.

Докажем второе утверждение леммы 4. Пусть $q = 1-p$. Из соотношений (30), (31) и усиленного закона больших чисел для неотрицательных ограниченных m-зависимых случайных величин (см. [23, теорема 1]) следует, что

$$\frac{1}{n}(T_{\mathrm{serial1}}(G) - T_{\mathrm{serial1}}(\underline{G})) = \sum_{\underline{B}\in\underline{G}} \frac{Y_n^2(\underline{B}1) + Y_n^2(\underline{B}0)}{n} - \sum_{\underline{B}\in\underline{G}} \left(\frac{2^{\frac{m-2}{2}}}{n} \big(\nu_{\underline{B}1} - \nu_{\underline{B}0}\big) \right)^2$$

$$= 2^{m-1} \sum_{\underline{B}\in\underline{G}} \left(\Big(\frac{\nu_{\underline{B}11} - \nu_{\underline{B}10}}{n}\Big)^2 + \Big(\frac{\nu_{\underline{B}01} - \nu_{\underline{B}00}}{n}\Big)^2 - \frac{1}{2}\Big(\frac{\nu_{\underline{B}1} - \nu_{\underline{B}0}}{n}\Big)^2 \right)$$

$$\xrightarrow{\text{п.н.}} 2^{m-1}\big((p_{\underline{B}11} - p_{\underline{B}10})^2 + (p_{\underline{B}01} - p_{\underline{B}00})^2 - \frac{1}{2}(p_{\underline{B}1} - p_{\underline{B}0})^2\big)$$

$$= 2^{m-1} p_B^2 \left((p^2 - pq)^2 + (pq - q^2)^2 - \frac{1}{2}(p-q)^2 \right) = 2^{m-1} p_B^2 (p-q)^2 \Big(p^2 + q^2 - \frac{1}{2}\Big) > 0.$$

Значит, $T_{\mathrm{serial2}} = T_{\mathrm{serial1}}(G) - T_{\mathrm{serial1}}(\underline{G}) \xrightarrow{\text{п.н.}} +\infty$. Лемма 4 доказана.

Автор выражает благодарность А. М. Зубкову за постоянное внимание к работе.

Список литературы

1. Rukhin A., Soto J., Nechvatal J., Smid M., Barker E., Leigh S., Levenson M., Vangel M., Banks D., Heckert A., Dray J., Vo S., "A statistical test suite for random and pseudorandom number generators for cryptographic applications", *NIST Special Publication* 800-22, *Revision* 1*a*, April 2010.
2. Серов А. А., "Формулы для чисел последовательностей, содержащих заданный шаблон заданное число раз", *Дискретная математика*, **32**:4 (2020), 120–136.
3. Zubkov A. M., Serov A. A., "A natural approach to the experimental study of dependence between statistical tests", *Матем. вопр. криптогр.*, **12**:1 (2021), 131–142.
4. Zubkov A. M., Serov A. A., "Testing the NIST Statistical Test Suite on artificial pseudorandom sequences", *Матем. вопр. криптогр.*, **10**:2 (2019), 89–96.
5. Zaman J. K. M. S. , Ghosh R., "Review on fifteen statistical tests proposed by NIST", *J. Theor. Phys. Cryptography*, **1** (2012), 18–31.
6. Sulak F., Doğanaksoy A., Uğuz M., Koçak O., "Periodic template tests: A family of statistical randomness tests for a collection of binary sequences", *Discrete Applied Mathematics*, **271** (2019), 191–204.
7. Soto J., Bassham L., "Randomness testing of the Advanced Encryption Standard finalist candidates", *NIST IR* 6483, *Nat. Inst. Stand. Technol.*, 2000.
8. Sulak F., Uğuz M., Koçak O., Doğanaksoy A., "On the independence of statistical randomness tests included in the NIST test suite", *Turkish J. Electr. Eng. & Comput. Sci.*, *T.*, **25**:5 (2017), 3673–3683.
9. Georgescu C., Simion E., "New results concerning the power of NIST randomness tests", *Proc. Romanian acad., ser. A.*, **18** (2017), 381–388.
10. Iwasaki A., Umeno K., "Randomness test to solve Discrete Fourier Transform Test problems", *IEICE Trans. Fundam. Electronics, Commun. Comput. Sci.*, **E101.A**:8 (2018), 1204–1214.
11. Burciu P., Simion E., "A systematic approach of NIST statistical tests dependencies", *J. Electr. Eng., Electronics, Control and Comput. Sci.*, **5**:15 (2019), 1–6.
12. Rukhin A. L., "Testing randomness: a suite of statistical procedures", *Teor. Veroyatnost. i Primenen.*, **45**:1 (2000), 137—162.
13. Рябко Б. Я., Пестунов А. И., "«Стопка книг» как новый статистический тест для случайных чисел", *Пробл. передачи информ.*, **40**:1 (2004), 73–78.
14. Мальцев М. В., Харин Ю. С., "О тестировании выходных последовательностей криптографических генераторов на основе цепей Маркова условного порядка", *Информатика*, **4** (2013), 104–111.
15. Савелов М. П., "Предельные совместные распределения статистик четырех критериев пакета NIST", *Дискретная математика*, **33**:2 (2021), 141–154.
16. Савелов М. П., "Предельные совместные распределения статистик трех критериев пакета NIST", *Дискретная математика*, **33**:3 (2021), 92–106.
17. Савелов М. П., "Предельная теорема для сглаженного варианта спектрального критерия равновероятности двоичной последовательности", *Дискретная математика*, **33**:4 (2021), 132–140.
18. Савелов М. П., "Предельное совместное распределение статистик критериев «Monobit test», «Frequency Test within a Block» и «Test for the Longest Run of Ones in a Block»", *Дискретная математика*, **34**:3 (2022), 70–84.
19. Савелов М. П., "Предельное совместное распределение статистик критериев «Monobit test», «Frequency Test within a Block» и «Binary Matrix Rank Test»", *Дискретная математика*, **34**:4 (2022), 84–98.
20. Good I. J., "The serial test for sampling numbers and other tests for randomness", *Proc. Camb. Phil. Soc.*, **49**:2 (1953), 276–284.

21. Good I. J., Gover T. N., “The generalized serial test and the binary expansion of $\sqrt{2}$”, *J. Roy. Statist. Soc. Ser. A*, **130**:1 (1967), 102–107.
22. Биллингсли П., *Сходимость вероятностных мер*, перев. с англ., «Наука», М., 1977, 352 pp.
23. Петров В. В., “Об усиленном законе больших чисел для последовательности неотрицательных случайных величин”, *Теория вероятн. и ее примен.*, **53**:2 (2008), 379–382.
24. Степанов В. Е., “Некоторые статистические критерии для цепей Маркова”, *Теория вероятн. и ее примен.*, **2**:1 (1957), 143–144.

Статья поступила 01.09.2022.