Дискретная математика

том 21 выпуск 3 * 2009

УДК 519.2

Об оценке периодов зашумленных двоичных периодических последовательностей

© 2009 г. М. И. Тихомирова, В. П. Чистяков

Для оценки длин периодов зашумленных двоичных последовательностей предлагаются и исследуются статистики, равные логарифму функции правдоподобия, и близкие к ним.

1. Введение

При проверке качества последовательностей псевдослучайных чисел $\{x_t^*\}$ за основную гипотезу H_0 принимается предположение, по которому исследуемая последовательность $\{x_t^*\}$ является реализацией последовательности независимых равномерно распределенных случайных величин $\{x_t\}$. Во множество предполагаемых конкурирующих гипотез включается обычно предположение о периодичности (в том или ином смысле) последовательности $\{x_t\}$.

Одним из возможных подходов к проверке предполагаемой периодичности является подход, основанный на использовании дискретного преобразования Фурье (см. [1, 2]). Пусть

$$x_1, x_2, \dots, x_t, \dots \tag{1}$$

— случайные величины, принимающие значения 0 и 1. Положим

$$Y_j = \frac{T}{2}|X(\omega_j)|^2, \quad j = 1, \dots, L, \quad L = [T/2], \tag{2}$$

где

$$X(\omega) = \frac{2}{T}\sum_{t=1}^{T}(2x_t - 1)e^{i\omega t}, \qquad \omega_j = \frac{2\pi j}{T}.$$

По гипотезе H_0 случайные величины (1) независимы и вероятности каждого из знаков 0 и 1 равны $1/2$. Для проверки гипотезы H_0 используем статистику $Y_{j_0} = \max_j Y_j$. Если гипотеза H_0 неверна, то в качестве оценок частоты ω_0 и соответствующего ей истинного периода n_0 принимаются $\widehat{\omega} = \omega_{j_0}$ и $\hat{n}_0 = T/j_0$.

В нашей работе исследуются двоичные последовательности (1), которые по конкурирующей гипотезе H_1 являются зашумленными случайными периодическими последовательностями, определяемыми следующим образом. Пусть $\{z_t\}$, $\{y_t\}$, $\{\varepsilon_t\}$ — случайные двоичные последовательности и

$$\mathbf{P}\{z_t = 1\} = \mathbf{P}\{y_t = 1\} = 1/2, \quad \mathbf{P}\{\varepsilon_t = 1\} = p, \quad 0 < p < 1, \quad t = 1, 2, \dots$$

Случайные величины $y_1, \ldots, y_{n_0}$, z_t, ε_t, $t = 1, 2, \ldots$, независимы. Последовательность $\{y_t\}$ является периодической:

$$\mathsf{P}\{y_{n_0 l+m} = y_m,\ l = 0, 1, \ldots,\ m = 1, 2, \ldots, n_0\} = 1. \tag{3}$$

Зашумленная периодическая последовательность $\{x_t\}$ определяется следующим образом:

$$x_t = \varepsilon_t y_t + (1 - \varepsilon_t) z_t, \qquad t = 1, 2, \ldots \tag{4}$$

В работе получено явное выражение для функции правдоподобия выборки, состоящей из T первых членов последовательности (1), предложены статистики, которые могут быть использованы для оценки периода последовательности, исследуются свойства этих статистик.

2. Функция правдоподобия

Напомним, что n_0 — период последовательности $\{y_t\}$. Найдем распределение первых $T = n_0 L_0$, $L_0 > 1$, членов последовательности $\{x_t\}$. Положим $x(T) = (x_1, \ldots, x_T)$, $y(n_0) = (y_1, \ldots, y_{n_0})$. Возможные значения этих векторов обозначим через $g(T)$ и $h(n_0)$:

$$g(T) = (g_1, \ldots, g_T), \quad h(n_0) = (h_1, \ldots, h_{n_0}), \quad g_t, h_t \in \{0, 1\}.$$

Теорема 1. *Совместное распределение $T = n_0 L_0$ первых членов последовательности $\{x_t\}$, определенной равенством* (4)*, имеет следующий вид:*

$$\mathsf{P}\{x(T) = g(T)\} = \frac{1}{2} \left(\frac{1-p}{2} \right)^{n_0 L_0} e^{G_{n_0}(g(T))}, \tag{5}$$

где

$$G_{n_0}(T) = \sum_{m=1}^{n_0} \ln(A^{\chi_m(n_0)} + A^{L_0 - \chi_m(n_0)}), \qquad A = \frac{1+p}{1-p},$$

$$\chi_m(n_0) = \sum_{l=0}^{L_0-1} \chi(g_{n_0 l+m} = 1), \qquad m = 1, \ldots, n_0,$$

$\chi(C)$ — индикатор события C.

Доказательство. Используя равенство (4) при $t = n_0 l + m$, получим, что

$$\begin{aligned}\mathsf{P}\{x(T) = g(T) \mid y(n_0) = h(n_0)\} &= \prod_{l=0}^{L_0-1} \prod_{m=l}^{n_0} \mathsf{P}\{\varepsilon_{n_0 l+m} h_m + (1 - \varepsilon_{n_0 l+m}) z_{n_0 l+m} = g_{n_0 l+m}\} \\ &= \prod_{l=0}^{L_0-1} \prod_{m=l}^{n_0} (p\chi(h_m = g_{n_0 l+m}) + (1-p)/2).\end{aligned}$$

Отсюда и из равенств

$$1 + \frac{2p}{1-p} \chi(h_m = g_{n_0 l+m}) = \left(1 + \frac{2p}{1-p}\right)^{\chi(h_m = g_{n_0 l+m})}, \qquad \mathsf{P}\{y(n_0) = h(n_0)\} = \frac{1}{2^{n_0}}$$

следует, что

$$\mathbf{P}\{x(T)=g(T),\, y(n_0)=h(n_0)\} = \frac{(1-p)^{n_0L_0}}{2^{n_0(L_0+1)}} A^{\sum_{m=1}^{n_0}\sum_{l=0}^{L_0-1}\chi(h_m=g_{n_0l+m})}. \tag{6}$$

Суммируя обе стороны равенства (6) по всем значениям $h(n_0)$, получим, что

$$\mathbf{P}\{x(T)=g(T)\} = \frac{(1-p)^{n_0L_0}}{2^{n_0(L_0+1)}} \sum_{h_1,\dots,h_{n_0}} A^{\sum_{m=1}^{n_0}\sum_{l=0}^{L_0-1}\chi(h_m=g_{n_0l+m})}.$$

Отсюда следует утверждение теоремы.

Если период n_0 неизвестен, то для оценки периода последовательности $\{y_t\}$ по последовательности $\{x_t\}$ можно использовать статистики

$$\xi_n(T) = \sum_{m=1}^{n} \ln\left(A^{\nu_m^{(T)}(n)} + A^{L_n-\nu_m^{(T)}(n)}\right), \qquad n = 2, 3, \dots, \tag{7}$$

где $L=[T/n]$, а частоты $\nu_m^{(T)}(n)$ определяются равенствами

$$\nu_m^{(T)}(n) = \sum_{l=0}^{L_n-1} x_{nl+m}. \tag{8}$$

Будет показано, что при больших T распределение статистики $\xi_{n_0}(T)$ заметно отличается от распределений статистики $\xi_n(T)$ с n, не кратным n_0.

3. Моменты и распределения частот

Будем считать, что неизвестный период n_0 последовательность $\{y_t\}$ является одним из $r+1$ натуральных чисел заданного множества. Числа этого множества, отличные от n_0, обозначим через $n_1, \dots, n_r$. Частоты (8) удобно записать в виде

$$\nu_m^{(T)}(n_i) = \sum_{t\in Z_m(n_i)} x_t, \qquad i = 0, 1, \dots, r, \tag{9}$$

где

$$Z_m(n_i) = \{t\colon t = n_i l + m,\ l = 0, 1, \dots, [T/n_i]-1\}, \qquad m = 1, 2, \dots, n_i. \tag{10}$$

Множества (10), соответствующие истинному периоду, обозначим $M_m = Z_m(n_0)$. Для упрощения формул ковариаций частот (9) ограничимся рассмотрением случая, когда

$$T = n_0 n_1 \dots n_r L, \tag{11}$$

где L — целое число, $L \geqslant 1$, числа n_i, $i = 0, 1, \dots, r$, являются попарно взаимно простыми. Число элементов множества C будем обозначать через $|C|$.

При условии (11)

$$|Z_m(n_i)| = \frac{T}{n_i}, \quad |Z_{m_1}(n_i) \cap Z_{m_2}(n_j)| = \frac{T}{n_i n_j}, \quad i \neq j. \tag{12}$$

Математические ожидания, дисперсии, ковариации и распределения вероятностей при гипотезах H_0 и H_1 будем обозначать $\mathbf{E}_l$, $\mathbf{D}_l$, cov_l, $\mathbf{P}_l$, $l = 0, 1$, соответственно.

При гипотезе H_0

$$\mathbf{E}_0 x_t = 1/2, \qquad \mathbf{D}_0 x_t = 1/4.$$

Отсюда, учитывая независимость слагаемых в сумме (9), получаем, что

$$\mathbf{E}_0 \nu_m^{(T)}(n_i) = \frac{T}{2n_i}, \qquad \mathbf{D}_0 \nu_m^{(T)}(n_i) = \frac{T}{4n_i}. \tag{13}$$

Для вычисления ковариаций частот (9) воспользуемся равенством

$$\mathrm{cov}_0(\nu_{m_1}^{(T)}(n_i), \nu_{m_2}^{(T)}(n_j)) = \sum_{t_1 \in Z_{m_1}(n_i)} \sum_{t_2 \in Z_{m_2}(n_j)} \mathrm{cov}_0(x_{t_1}, x_{t_2}) = \sum_{t \in Z_{m_1}(n_i) \cap Z_{m_2}(n_j)} \mathbf{D}_0 x_t.$$

Отсюда и из равенств (12) и (13) получаем, что

$$\begin{aligned} &\mathrm{cov}_0(\nu_{m_1}^{(T)}(n_i), \nu_{m_2}^{(T)}(n_j)) = \frac{T}{4n_i n_j}, && i \neq j, \\ &\mathrm{cov}_0(\nu_{m_1}^{(T)}(n_i), \nu_{m_2}^{(T)}(n_j)) = 0, && i = j, \quad m_1 \neq m_2. \end{aligned} \tag{14}$$

Используя многомерную центральную предельную теорему и формулы (13) и (14), получим следующее утверждение.

Теорема 2. *Если $n_0, n_1, \ldots, n_r$ постоянны, взаимно просты, и $L \to \infty$, то при гипотезе H_0 совместное распределение величин*

$$\theta_m^{(T)}(n_i) = \frac{\nu_m^{(T)}(n_i) - T/(2n_i)}{\sqrt{T/n_i}/2}, \quad m = 1, \ldots, n_i, \quad i = 0, 1, \ldots, r,$$

сходится к распределению случайного вектора $\{\theta_m^*(n_i)\}$, имеющего нормальное распределение с нулевым вектором математических ожиданий и ковариационной матрицей $(\mathrm{cov}_0(\nu_{m_1}^*(n_i), \nu_{m_2}^*(n_j)))$, где

$$\begin{aligned} &\mathrm{cov}_0(\nu_{m_1}^*(n_i), \nu_{m_2}^*(n_j)) = 1, && i = j, \quad m_1 = m_2, \\ &\mathrm{cov}_0(\nu_{m_1}^*(n_i), \nu_{m_2}^*(n_j)) = 0, && i = j, \quad m_1 \neq m_2, \\ &\mathrm{cov}_0(\nu_{m_1}^*(n_i), \nu_{m_2}^*(n_j)) = \sqrt{n_i n_j}, && i \neq j. \end{aligned} \tag{15}$$

При гипотезе H_1 для моментов величин x_t, определенных равенством (4), нетрудно получить следующие выражения:

$$\begin{aligned} &\mathbf{E}_1 x_t = 1/2, && && \mathbf{D}_1 x_t = 1/4, \\ &\mathrm{cov}_1(x_{t_1}, x_{t_2}) = 0, && t_1 \in M_u, \quad t_2 \in M_v, && u \neq v, \\ &\mathrm{cov}_1(x_{t_1}, x_{t_2}) = p^2/4, && t_1, t_2 \in M_u, && t_1 \neq t_2. \end{aligned} \tag{16}$$

Отсюда следует, что математическое ожидание частоты $\nu_m^{(T)}(n_i)$ при гипотезе H_1 совпадает с соответствующим математическим ожиданием при гипотезе H_0, то есть

$$\mathbf{E}_1 \nu_m^{(T)}(n_i) = \frac{T}{2n_i}. \tag{17}$$

Частоты (9) представимы в виде

$$\nu_m^{(T)}(n_i) = \sum_{u=1}^{n_0} \left(\sum_{t \in Z_m(n_i) \cap M_u} x_t \right), \qquad \nu_m^{(T)}(n_0) = \sum_{t \in M_m} x_t, \tag{18}$$

где суммы в круглых скобках независимы.

Используя представление частот в виде (18), нетрудно получить следующее утверждение.

Теорема 3. *Если последовательность* $\{x_t\}$ *определяется равенствами* (3) *и* (4) *и* T *имеет вид* (11)*, то*

$$\mathbf{D}_1 \nu_m^{(T)}(n_i) = \frac{1}{4}\left((1-p^2)\frac{T}{n_0} + \frac{T^2}{n_0^2} p^2 \right), \qquad i = 0,$$

$$\mathbf{D}_1 \nu_m^{(T)}(n_i) = \frac{1}{4}\left((1-p^2)\frac{T}{n_i} + \frac{T^2}{n_0 n_i^2} p^2 \right), \qquad i \neq 0,$$

$$\operatorname{cov}_1(\nu_{m_1}^{(T)}(n_i), \nu_{m_2}^{(T)}(n_j)) = \frac{1}{4}\left((1-p^2)\frac{T}{n_i n_j} + \frac{T^2}{n_i n_j n_0} p^2 \right), \qquad i \neq j, \quad i \neq 0, \quad j \neq 0,$$

$$\operatorname{cov}_1(\nu_{m_1}^{(T)}(n_i), \nu_{m_2}^{(T)}(n_j)) = \frac{1}{4}\left((1-p^2)\frac{T}{n_0 n_j} + \frac{T^2}{n_0^2 n_j} p^2 \right), \qquad i = 0, \quad j \neq 0,$$

$$\operatorname{cov}_1(\nu_{m_1}^{(T)}(n_i), \nu_{m_2}^{(T)}(n_j)) = \frac{1}{4}\left(p^2 \frac{T^2}{n_i^2 n_0^2} \right), \qquad i = j \neq 0, \quad m_1 \neq m_2,$$

$$\operatorname{cov}_1(\nu_{m_1}^{(T)}(n_i), \nu_{m_2}^{(T)}(n_j)) = 0, \qquad i = j = 0, \quad m_1 \neq m_2.$$

Докажем, например, формулу для случая $i \neq j$, $i \neq 0$, $j \neq 0$. Из представления $\nu_m^{(T)}(n_i)$ в виде (18), учитывая независимость сумм в круглых скобках, получаем равенство

$$\operatorname{cov}_1(\nu_{m_1}(n_i), \nu_{m_2}(n_j)) = \sum_{u=1}^{n_0} \sum_{t_1 \in Z_{m_1}(n_i) \cap M_u} \sum_{t_2 \in Z_{m_2}(n_j) \cap M_u} \operatorname{cov}_1(x_{t_1}, x_{t_2}). \tag{19}$$

Сумму по t_1, t_2 представим в виде суммы двух слагаемых

$$\Sigma_1(u) = \sum_{t \in Z^{(1)}} \mathbf{D}_1 x_t, \qquad \Sigma_2(u) = \sum_{(t_1,t_2) \in Z^{(2)}} \operatorname{cov}_1(x_{t_1}, x_{t_2}),$$

где $Z^{(1)} = Z_{m_1}(n_i) \cap Z_{m_2}(n_j) \cap M_u$, а $Z^{(2)}$ состоит из пар (t_1, t_2) с $t_1 \neq t_2$. Нетрудно проверить, что

$$|Z^{(1)}| = \frac{T}{n_i n_j n_0}, \qquad |Z^{(2)}| = \frac{T}{n_i n_0}\frac{T}{n_j n_0} - \frac{T}{n_i n_j n_0}. \tag{20}$$

Из (16), (18) и (20) получаем первую формулу для ковариаций. Остальные формулы доказываются аналогично.

По гипотезе H_1, дисперсии частот при $T \to \infty$ имеют порядок T^2. Дисперсии частот по условному распределению при фиксированном векторе $y(n_0) = (y_1, y_2, \ldots, y_{n_0})$ имеют обычный порядок T. При $t \in M_m$

$$\mathbf{P}_1\{x_t = 1 \mid y_m = 1\} = \frac{1+p}{2}, \qquad \mathbf{P}_1\{x_t = 1 \mid y_m = 0\} = \frac{1-p}{2}. \tag{21}$$

При условии y_m величины $x_t, t \in M_m$ независимы и

$$\mathbf{E}_1(\nu_m^{(T)}(n_0) \mid y_m = 1) = \frac{T(1+p)}{2n_0},$$
$$\mathbf{E}_1(\nu_m^{(T)}(n_0) \mid y_m = 0) = \frac{T(1-p)}{2n_0},$$
$$\mathbf{D}_1(\nu_m^{(T)}(n_0) \mid y_m = 1) = \mathbf{D}_1(\nu_m^{(T)}(n_0) \mid y_m = 0) = \frac{T(1-p^2)}{4n_0}. \tag{22}$$

Таким образом, при $T \to \infty$

$$\mathbf{P}_1\left\{\frac{\nu_m^{(T)}(n_0) - T(1+p)/(2n_0)}{(1/2)\sqrt{T(1-p^2)/n_0}} < x \mid y_m = 1\right\} \to \Phi(x),$$
$$\mathbf{P}_1\left\{\frac{\nu_m^{(T)}(n_0) - T(1-p)/(2n_0)}{(1/2)\sqrt{T(1-p^2)/n_0}} < x \mid y_m = 0\right\} \to \Phi(x),$$

где $\Phi(x)$ — функция нормального распределения с параметрами $(0, 1)$. Таким образом, распределение $\nu_m^{(T)}(n_0)$ сосредоточено с равными вероятностями вокруг удаляющихся друг от друга точек $T(1+p)/(2n_0)$ и $T(1-p)/(2n_0)$. Распределение $\nu_m(n_i)$ при $i \neq 0$ является смесью $n_0 + 1$ распределений, зависящих от суммы $\mu_{n_0} = y_1 + \ldots + y_{n_0}$.

4. Моменты и распределения статистик

Математические ожидания частот $\nu_m^{(T)}(n_i)$ при обеих гипотезах одинаковы. Статистики (7), зависящие от частот $\nu_m^{(T)}(n_i)$, можно заменить статистиками, зависящими от центрированных частот и отличающимися от частот (7) постоянными слагаемыми:

$$\xi_n^*(T) = \sum_{m=1}^{n} \ln(A^{\hat\theta_m^{(T)}(n)} + A^{-\hat\theta_m^{(T)}(n)}), \qquad n = n_0, n_1, \ldots, n_r, \tag{23}$$

где

$$\hat\theta_m^{(T)}(n) = \nu_m^{(T)}(n) - \frac{T}{2n}, \qquad A = \frac{1+p}{1-p}. \tag{24}$$

Теорема 4. *Если $n_0, n_1, \ldots, n_r$ постоянны, взаимно просты, $T = n_0 n_1 \ldots n_r L$ и $L \to \infty$, то при гипотезе H_0 распределение вектора $(2/\sqrt{A}\ln A)(\sqrt{n_0}\xi_{n_0}^*(T), \ldots, \sqrt{n_r}\xi_{n_r}^*(T))$ сходится к совместному распределению величин $\sum_{m=1}^{n_i} |\theta_m^*(n_i)|$, $i = 0, 1, \ldots, r$, где величины $\theta_m^*(n_i)$ определены в формулировке теоремы 2.*

Доказательство. Представим статистики (23) в виде

$$\xi_n^*(T) = \sum_{m=1}^{n} (\ln A^{|\theta_m^{(T)}(n)|\sqrt{T/n}/2} + \Delta_m^{(T)}(n)), \tag{25}$$

где

$$\theta_m^{(T)}(n) = 2\hat\theta_m^{(T)}(n)/\sqrt{T/n},$$
$$\Delta_m^{(T)}(n) = \ln(A^{(\theta_m^{(T)}(n) - |\theta_m^{(T)}(n)|)\sqrt{T/n}/2} + A^{-(\theta_m^{(T)}(n) - |\theta_m^{(T)}(n)|)\sqrt{T/n}/2}). \tag{26}$$

Положим $\gamma_m = 1$, если $\theta_m^{(T)}(n) > 0$, и $\gamma_m = 0$, если $\theta_m^{(T)} Z(n) < 0$.

Второе равенство (26) можно записать в виде

$$\Delta_m^{(T)}(n) = \gamma_m \ln(1 + A^{-\theta_m^{(T)}(n)\sqrt{T/n}}) + (1-\gamma_m)\ln(1 + A^{\theta_m^{(T)}(n)\sqrt{T/n}}).$$

Отсюда

$$\Delta_m^{(T)}(n) = \ln(1 + A^{-|\theta_m^{(T)}(n)|\sqrt{T/n}})$$

Из теоремы 2 следует, что величины $\Delta_m^{(T)}(n)$ при $L \to \infty$ сходятся по вероятности к нулю. Таким образом, предельные распределения $\xi_{n_i}^*(T)$ и $\sum_{m=1}^{n_i}\left(\frac{1}{2}\sqrt{T/n_i}\ln A\right)|\theta_m^{(T)}(n_i)|$, $i = 0, 1, \ldots, r$, совпадают. Теорема доказана.

Рассмотрим более простые (по сравнению с (23)) статистики, для которых можно получать простые явные выражения для моментов. Разлагая слагаемые в (23) в ряд Тейлора по степеням $\hat{\theta}_n^{(T)}(n)$, получим, что

$$\ln(A^{\hat{\theta}_m^{(T)}(n)} + A^{-\hat{\theta}_m^{(T)}(n)}) = \ln 2 + \frac{1}{2}(\hat{\theta}_m^{(T)}(n))^2(\ln A)^2 + \ldots$$

Определим статистики $\eta_{n_i}(T)$ равенствами

$$\eta_n(T) = \sum_{m=1}^{n}(\hat{\theta}_m^{(T)}(n))^2, \qquad n = n_0, n_1, \ldots, n_r, \tag{27}$$

где $\hat{\theta}_m^{(T)}(n)$ определены равенствами (24).

При гипотезах H_0 и H_1 из формул (13), (17) и (27) следует, что

$$\mathbf{E}_l \eta_n(T) = \sum_{m=1}^{n} \mathbf{D}_l \nu_m(n), \qquad l = 0, 1.$$

Отсюда, из формул (13) и теоремы 3 получаем следующие два утверждения.

Следствие 1. *При гипотезе* H_0 *и* $T = n_0 n_1 \ldots n_r L$

$$\mathbf{E}_0 \eta_n(T) = T/4, \qquad n = n_0 n_1 \ldots n_r.$$

Следствие 2. *Если числа* $n_0, n_1, \ldots, n_r$ *взаимно просты, то при гипотезе* H_1 *и* $T = n_0 n_1 \ldots n_r L$

$$\mathbf{E}_1 \eta_{n_0}(T) = \frac{1}{4}(p^2 T^2/n_0 + (1-p^2)T),$$
$$\mathbf{E}_1 \eta_{n_0}(T) = \frac{1}{4}(p^2 T^2/(n n_0) + (1-p^2)T), \qquad n = n_0, \ldots, n_r.$$

Из полученных формул следует, что

$$\lim_{T\to\infty} \frac{4n_0}{p^2T^2}\mathbf{E}_1 \eta_n(T) = 1, \qquad n = n_0,$$
$$\lim_{T\to\infty} \frac{4n_0}{p^2T^2}\mathbf{E}_1 \eta_n(T) = \frac{1}{n_i}, \qquad n = n_1, \ldots, n_r,$$

Если получить несколько реализаций T первых членов последовательности $\{x_t\}$, то можно оценить $\mathbf{E}_1 \eta_n(T)$ и найти n_0. Возможность определения n_0 по одной реализации рассматривается в следующем параграфе.

5. Моменты статистик относительно условных распределений

При любом $T > n_0$ случайные величины x_t, определенные равенствами (4), зависят от одних и тех же величин $y(n_0) = (y_1, \ldots, y_{n_0})$. Найдем моменты частот и статистик при заданных значениях $y(n_0)$.

Из формул (21) и (22) следует, что

$$\mathbf{P}_1\{x_t = 1 \mid y_m\} = \frac{1}{2}(1 - (1 - 2y_m)p), \qquad t \in M_m. \tag{28}$$

Величины $\hat{\theta}_m^{(T)}(n)$ представим в виде

$$\hat{\theta}_m^{(T)}(n_0) = \sum_{t \in M_n} (x_t - 1/2),$$
$$\hat{\theta}_m^{(T)}(n) = \sum_{u=1}^{n_0} \sum_{t \in Z_m(n) \cap M_u} (x_t - 1/2), \qquad n = n_1, \ldots, n_r. \tag{29}$$

Из равенства (28) при $t \in M_u$ получаем, что $\mathbf{E}_1(x_t - 1/2) \mid y_m) = p(y_m - 1/2)$. Отсюда следует, что по условному распределению с фиксированным $y(n_0)$ средние арифметические одинаково распределенных случайных величин $(n_i n_0/T) \sum_{t \in Z_m(n_i) \cap M_u} (x_t - 1/2)$, $(n_0/T) \sum_{t \in M_m} (x_t - 1/2)$ при $T \to \infty$ сходится по вероятности к величинам $p(y_u - 1/2)$ и $p(y_m - 1/2)$ соответственно. Используя (29), получаем, что величины $(1/T)\hat{\theta}_m^{(T)}(n_i)$, $(1/T)\hat{\theta}_m^{(T)}(n_0)$ сходятся по вероятности к $(p/(n_i n_0))(\mu_{n_0} - n_0/2)$, $(p/n_0)(y_m - 1/2)$, $\mu_{n_0} = y_1 + \ldots + y_{n_0}$, а также сходятся случайные величины $\eta_{n_i}(T)/T^2$ и $\eta_{n_0}(T)/T^2$. Таким образом, доказано следующее утверждение.

Теорема 5. *Совместное условное распределение случайных величин $\eta_{n_i}(T)/T^2$, $\eta_0(T)/T^2$, $i = 1, \ldots, r$, с фиксированным $y(n_0) = (y_1, \ldots, y_{n_0})$ при $T \to \infty$ слабо сходится к вырожденному распределению, сосредоточенному в точке*

$$\frac{p^2}{4n_0}, \quad \frac{p^2}{n_i n_0^2}\left(\mu_{n_0} - \frac{n_0}{2}\right)^2, \quad i = 1, \ldots, r. \tag{30}$$

6. Примеры

Рассмотрим статистики

$$\zeta_n^{(1)}(T) = \xi_n^*(T), \quad \zeta_n^{(2)}(T) = \eta_n(T), \quad \zeta_m^{(3)}(T) = Y_n^*(T), \tag{31}$$

где $\zeta_n^{(1)}(T)$, $\zeta_n^{(2)}(T)$ определены формулами (23), (24) и (27), а

$$Y_n^*(T) = \frac{T}{2}|X(2\pi/n)|^2, \qquad X(\omega) = \frac{2}{T}\sum_{t=1}^{T}(2x_t - 1)\exp\{i\omega t\}. \tag{32}$$

Статистики (31) введены вместо статистик (2) для удобства сравнения со статистиками $\zeta_n^{(k)}(T)$. Статистики (2) зависят от частот $\omega_j = 2\pi j/T$ и занумерованы числами j, а статистики (32) занумерованы периодами n, соответствующими частотам ω_j с $j = T/n$.

Таблица 1.

k	$\zeta_n^{(k)}(T)$	$\widehat{E}_{12}(k)$	$\widehat{E}_{13}(k)$	$\widehat{E}_{14}(k)$	$\hat{p}_{12}(k)$	$\hat{p}_{13}(k)$	$\hat{p}_{14}(k)$
1	$\xi_n^*(T)$	22,1	46,1	23,5	0	0,993	0,002
2	$\eta_n(T)$	113	413	105	0	0,983	0
3	$Y_n^*(T)$	2,01	8,59	1,81	0,049	0,590	0,049

Таблица 2.

k	$\zeta_n^{(k)}(T)$	$\widehat{E}_{12}(k)$	$\widehat{E}_{13}(k)$	$\widehat{E}_{14}(k)$	$\hat{p}_{12}(k)$	$\hat{p}_{13}(k)$	$\hat{p}_{14}(k)$
1	$\xi_n^*(T)$	12,6	16,1	14,0	0,002	0,484	0,054
2	$\eta_n(T)$	99,3	179	96,6	0,037	0,620	0,035
3	$Y_n^*(T)$	1,96	3,70	2,00	0,083	0,270	0,095

Пусть $n_0 = 13$ и возможное множество периодов $\{n_0, n_1, \dots, n_8\} = \{8, 9, \dots, 16\}$, $T = 390$. Оценку $\hat{n}_0(k)$ периода n, найденную по последовательности (1) с использованием статистики $\zeta_n^{(k)}(T)$, определим равенством

$$\zeta_{\hat{n}_0(k)}^{(k)}(T) = \max_{8 \leqslant n \leqslant 16} \zeta_n^{(k)}(T),$$

Для иллюстрации свойств статистик (31) было смоделировано по 1000 реализаций $T = 390$ первых членов последовательности (1) при гипотезе H_1 с $p = 1/3$ и $p = 1/6$. Получены оценки $\widehat{E}_n(k)$, $\hat{p}_n(k)$ величин

$$\mathbf{E}\zeta_n^{(k)}(T) = E_n(k), \quad p_n(k) = \mathbf{P}\{\hat{n}_0(k) = n\}, \quad k = 1, 2, 3, \quad 12 \leqslant n \leqslant 14.$$

Эти оценки приведены в табл. 1 для $p = 1/3$ и табл. 2 для $p = 1/6$.

Список литературы

1. Бриллинджер Д., *Временные ряды*. Мир, Москва, 1980.

2. Харин Ю. С., Берник В. И., Матвеев Г. В., Агиевич С. В., *Математические и компьютерные основы криптографии*. Новое знание, Минск, 2003.

Статья поступила 14.09.2006.