Дискретная математика

том 19 выпуск 2 * 2007

УДК 519.2

Цепь Маркова s-го порядка с r частичными связями и статистические выводы о ее параметрах

© 2007 г. Ю. С. Харин, А. И. Петлицкий

Рассматривается однородная цепь Маркова s-го порядка с r частичными связями, для которой вероятность перехода процесса в будущее состояние зависит не от всех s предыдущих состояний, а лишь от r избранных состояний. Построены статистические оценки параметров и статистические критерии проверки гипотез о значениях параметров этой модели, установлены их асимптотические свойства. Предложены алгоритмы вычисления оценок параметров. Приводятся результаты компьютерных экспериментов.

1. Введение

При математическом моделировании сложных систем и процессов в экономике, технике, медицине, социологии, генетике и других приложениях часто возникает необходимость построения адекватных вероятностно-статистических моделей дискретных временных рядов $x_t \in A, t \in \mathbf{N}$, где пространство состояний $A = \{0, 1, \ldots, N-1\}$ – конечное множество мощности $N \geqslant 2$ с длинной памятью [16]. Известной моделью таких дискретных временных рядов является цепь Маркова достаточно высокого порядка $s \in \mathbf{N}$, определяющего длину памяти; если $s = 1$, то цепь Маркова называется простой, если $s > 1$ – сложной [7]. Однако для такой модели число параметров D растет экспоненциально при увеличении порядка s: $D = N^s(N-1)$, и для статистического оценивания параметров требуется иметь реализацию $x_1, \ldots, x_n$ не всегда доступной на практике длительности $n > D$. В связи с этим актуальна проблема построения малопараметрических моделей цепей Маркова высокого порядка. В [8] предложено использовать модель смесей простых цепей Маркова. В данной работе исследуется новая общая малопараметрическая модель цепи Маркова s-го порядка с r частичными связями ЦМ(s, r), предложенная в [9], строятся статистические оценки параметров, критерии проверки гипотез о значениях параметров и анализируются их вероятностные свойства. В [10] для частного случая ЦМ$(s, 2)$ этой модели решена задача статистической проверки гипотез о чистой случайности наблюдаемой случайной последовательности.

2. Цепь Маркова ЦМ(s, r)

Пусть $(\Omega, F, \mathbf{P})$ — основное вероятностное пространство; $x_t \in A$, $t \in \mathbf{N}$, — цепь Маркова s-го порядка; $r \in \{1, \dots, s\}$ — параметр, который будем называть числом связей; $M_r^0 = (m_1^0, \dots, m_r^0) \in \mathbf{M}$ — произвольный целочисленный r-вектор с упорядоченными компонентами $1 = m_1^0 < m_2^0 < \dots < m_r^0 \leqslant s$, который будем называть шаблоном (связей), $\mathbf{M}$ — множество всевозможных таких векторов с r компонентами, имеющее мощность $\mathbf{K} = |\mathbf{M}| = \binom{s-1}{r-1}$; $P = (p_{i_1,\dots,i_s,i_{s+1}})$, $i_1, \dots, i_{s+1} \in A$, — $(s+1) \times (s+1)$ матрица вероятностей одношаговых переходов цепи Маркова x_t

$$p_{i_1,\dots,i_s,i_{s+1}} = \mathbf{P}\{x_t = i_{s+1} \mid x_{t-1} = i_s, \dots, x_{t-s} = i_1\}, \qquad t > s,$$

$Q^0 = (q^0_{j_1,\dots,j_r,j_{r+1}})$, $j_1, \dots, j_{r+1} \in A$, — некоторая $(r+1) \times (r+1)$ стохастическая матрица.

Определение 1. Цепь Маркова $x_t \in A$ назовем цепью Маркова s-го порядка с r частичными связями и будем обозначать ЦМ(s, r), если ее вероятности одношаговых переходов имеют вид

$$p_{i_1,\dots,i_s,i_{s+1}} = q^0_{i_{m_1^0},\dots,i_{m_r^0},i_{s+1}}, \quad i_1, \dots, i_{s+1} \in A, \quad \{i_{m_1^0}, \dots, i_{m_r^0}\} \subseteq \{i_1, \dots, i_s\}. \tag{1}$$

Соотношение (1) означает, что вероятность перехода процесса в состояние i_{s+1} в момент времени t зависит не от всех s предыдущих состояний $i_1, \dots, i_s$ процесса, а лишь от r избранных состояний $i_{m_1^0}, \dots, i_{m_r^0}$. Таким образом, вместо $D = N^s(N-1)$ параметров матрица P вероятностей переходов ЦМ(s, r) полностью определяется $d = N^r(N-1) + r - 1$ параметрами Q^0 и M_r^0. Например, если $N = 2$, $s = 32$, $r = 3$, то $D \approx 4{,}1 \cdot 10^9$, $d = 10$.

Если Q^0, M_r^0 не зависят от времени t, то имеем однородную цепь Маркова s-го порядка с r-частичными связями, в противном случае неоднородную. Далее в этой статье мы рассматриваем однородную цепь ЦМ(s, r), удовлетворяющую условию эргодичности.

Заметим, что если $s = r$, $M_r^0 = (1, \dots, s)$, то $P = Q^0$ и ЦМ(s, s) есть цепь Маркова s-го порядка [7]. Конструктивным примером ЦМ(s, r) является бинарная $(N = 2)$ авторегрессия [11] s-го порядка с r ненулевыми коэффициентами, частным случаем которой является линейная рекуррентная последовательность над кольцом Z_2 [12], порожденная многочленом степени s с r ненулевыми коэффициентами.

3. Статистическое оценивание параметров ЦМ(s, r)

Рассмотрим задачу построения оценок максимального правдоподобия (ОМП) параметров $\widehat{M}_r$, шаблона $M_r^0 \in \mathbf{M}$ и $\widehat{Q}$ стохастической матрицы Q^0 по наблюдаемой реализации $X_n = (x_1, \dots, x_n) \in A^n$ длительности $n > s$. Введем обозначения, пусть $J_s = (j_1, \dots, j_s) = (J_{s-1}, j_s) \in A^s$ — мультииндекс s-го порядка; $\delta_{J_s, J_s'}$ — символ Кронекера для мультииндексов J_s, J_s'; $S_t(X_n; M_r) = (x_{t+m_1-1}, \dots, x_{t+m_r-1}) \in A^r$ — функция $A^n \times \mathbf{M} \to A^r$, которую условимся называть селектором r-го порядка с параметрами $M_r \in \mathbf{M}$ и $t \in \{1, \dots, n-s+1\}$; $\mathbf{I}\{B\} \in \{0, 1\}$ — индикатор события B; $\Pi_{K_s} = \mathbf{P}\{X_s = K_s\}$ — начальное s-мерное распределение вероятностей ЦМ(s, r);

$$\nu_{r+1}(J_{r+1}; M_r) = \sum_{t=1}^{n-s} \mathbf{I}\{S_t(X_n; M_{r+1}) = J_{r+1}\} \tag{2}$$

— частота $(r+1)$-граммы $J_{r+1} \in A^{r+1}$ для шаблона $M_{r+1} = (M_r, s+1)$, удовлетворяющая условию нормировки

$$\sum_{J_{r+1} \in A^{r+1}} \nu_{r+1}(J_{r+1} : M_r) \equiv n - s.$$

Условимся полагать, что если вместо какого-то индекса стоит точка, то это означает суммирование по всем возможным значениям этого индекса:

$$\nu_{r+1}(J_r \cdot : M_r) = \sum_{j_{r+1} \in A} \nu_{r+1}(J_{r+1} : M_r),$$

$$\nu_{r+1}(\cdot\, j_{r+1} : M_r) = \sum_{J_r \in A^r} \nu_{r+1}(J_{r+1} : M_r)$$

— накопленные статистики, отличающиеся не более чем на s от частоты r-граммы $J_r \in A^r$ и частоты символа $j_{r+1} \in A$ соответственно.

Исследуем вначале вопрос эргодичности ЦМ(s, r).

Теорема 1. *Цепь Маркова s-го порядка с r частичными связями, определяемая вероятностями перехода* (1), *является эргодической тогда и только тогда, когда найдется целое число $l \geqslant 0$ такое, что*

$$\min_{J_s, J_s' \in A^s} \sum_{K_l \in A^l} \prod_{i=1}^{s+l} q^0_{S_i((J_s, K_l, J_s') : M^0_{r+1})} > 0.$$

Доказательство. Понимая эргодичность цепи Маркова s-го порядка как эргодичность эквивалентной ей s-векторной цепи Маркова первого порядка с расширенным пространством состояний [7] $X_{(t)} = (x_{t-s+1}, \ldots, x_{t-1}, x_t)^T \in A^s$, $t \geqslant s$, можно записать необходимое и достаточное условие эргодичности: существует $m < \infty$ такое, что $\min_{J_s, J_s' \in A^s} p^{*(m)}_{J_s, J_s'} > 0$, где $p^{*(m)}_{J_s, J_s'}$ — вероятность перехода $X_{(t)}$ за m шагов из состояния J_s в J_s', $p^*_{J_s, J_s'} = \delta_{J_s^-, J'_{s-1}} p_{J_s, j_s'}$, $J_s^- = (j_2, \ldots, j_s) \in A^{s-1}$, $p_{J_s, j_s'} = q^0_{S_1(J_s : M_r^0), j_s'}$. Проводя некоторые упрощения, при $m = s + l$ получаем, что

$$(P^*)^{(s+l)}_{J_s, J_s'} = \sum_{K_l \in A^l} \prod_{i=1}^{s+l} p_{S_i((J_s, K_l, J_s') : M_{s+1})}, \qquad M_{s+1} = (1, 2, \ldots, s+1),$$

что приводит к требуемому результату.

Следствие 1. *Если*

$$\min_{J_{r+1} \in A^{r+1}} q^0_{J_{r+1}} > 0,$$

то ЦМ(s, r) *является эргодической.*

Лемма 1. *Для модели ЦМ(s, r) логарифмическая функция правдоподобия имеет вид*

$$l(Q, M_r) = \ln \Pi_{X_s} + \sum_{J_{r+1} \in A^{r+1}} \nu_{r+1}(J_{r+1} : M_r) \ln q_{J_{r+1}}.$$

Теорема 2. *Если шаблон $M_r \in \mathbf{M}$ фиксирован, то условная ОМП для матрицы Q^0 имеет вид ($J_{r+1} \in A^{r+1}$):*

$$\hat{Q} = \hat{Q}(M_r) = (\hat{q}_{J_{r+1}}(M_r)),$$

$$\hat{q}_{J_{r+1}}(M_r) = \begin{cases} \nu_{r+1}(J_{r+1} : M_r)/\nu_{r+1}(J_r \cdot : M_r), & \text{если } \nu_{r+1}(J_r \cdot : M_r) > 0, \\ 1/N, & \text{если } \nu_{r+1}(J_r \cdot : M_r) = 0. \end{cases} \quad (3)$$

Доказательство. ОМП (3) получается прямым решением задачи на условный экстремум

$$l(Q, M_r) \to \max_Q, \qquad \sum_{j_{r+1} \in A} q_{J_r, j_{r+1}} = 1 \quad (4)$$

с использованием выражения функции правдоподобия из леммы 1.

Пусть $\Pi^*_{K_s}$, $K_s \in A^s$, — стационарное распределение вероятностей эргодической ЦМ(s, r). В предположении ее стационарности ($\Pi_{K_s} = \Pi^*_{K_s}$, $K_s \in A^s$) (см. [13]) введем следующие обозначения. Пусть

$$\mu_{r+1}(J_{r+1} : M_r, M_r^0) = \mathbf{P}\{S_t(X_n : M_{r+1}) = J_{r+1}\}$$
$$= \sum_{K_{s+1} \in A^{s+1}} \mathbf{I}\{S_1(K_{s+1} : M_{r+1}) = J_{r+1}\} \Pi^*_{K_s} p_{K_{s+1}}, \quad J_{r+1} \in A^{r+1}, \quad (5)$$

— распределение вероятностей $(r+1)$-граммы для шаблона $M_{r+1} = (M_r, s+1)$;

$$\hat{\mu}_{r+1}(J_{r+1} : M_r) = \nu_{r+1}(J_{r+1} : M_r/(n-s)$$

— частотная оценка вероятности $\mu_{r+1}(J_{r+1} : M_r, M_r^0)$, которая с учетом (2) и свойств частот для цепей Маркова [14] является несмещенной, состоятельной оценкой;

$$H_{r+1}(M_r, M_r^0) = -\sum_{J_{r+1} \in A^{r+1}} \mu_{r+1}(J_{r+1} : M_r, M_r^0) \ln \mu_{r+1}(J_{r+1} : M_r, M_r^0) \geqslant 0 \quad (6)$$

— энтропия $(r+1)$-мерного распределения вероятностей (5);

$$I_{r+1}(M_r, M_r^0) = \sum_{J_{r+1} \in A^{r+1}} \mu_{r+1}(J_{r+1} : M_r, M_r^0) \times \ln \frac{\mu_{r+1}(J_{r+1} : M_r, M_r^0)}{\mu_{r+1}(J_r \cdot : M_r, M_r^0)\, \mu_{r+1}(\cdot j_{r+1} : M_r, M_r^0)} \geqslant 0 \quad (7)$$

— количество информации по Шеннону, содержащейся в r-грамме $S_t(X_n : M_r) \in A^r$ о будущем символе $x_{t+s} \in A$, т.е. средняя условная энтропия x_{t+s} при условии $S_t(X_n, M_r)$; $\hat{H}_{r+1}(M_r)$, $\hat{I}_{r+1}(M_r)$ — подстановочные оценки энтропии и количества информации, получающиеся подстановкой в (6), (7) вместо истинных вероятностей $\{\mu_{r+1}(J_{r+1} : M_r, M_r^0)\}$ их оценок $\{\hat{\mu}_{r+1}(J_{r+1} : M_r)\}$.

Теорема 3. *ОМП шаблона M_r^0 выражается через информационный критерий*

$$\widehat{M}_r = \arg \max_{M_r \in \mathbf{M}} \hat{I}_{r+1}(M_r), \quad (8)$$

причем, если целевая функция имеет не один глобальный максимум, то оценка $\widehat{M}_r$ задается в виде множества эквивалентных шаблонов.

Доказательство. В силу леммы 1, теоремы 2 и обозначений (5), (6) получаем задачу максимизации (4) логарифмической функции правдоподобия в виде

$$l_1(M_r) = l(\widehat{Q}(M_r), M_r) = \ln \Pi_{X_s} + (n-s)(\widehat{H}_r(M_r) - \widehat{H}_{r+1}(M_r)) \to \max_{M_r}.$$

Учитывая, что первый член не зависит от M_r, добавляя не зависящее от M_r слагаемое $(n-s)\widehat{H}_1$, используя (7) и тождество

$$\hat{I}_{r+1}(M_r) \equiv \widehat{H}_r(M_r) + \widehat{H}_1 - \widehat{H}_{r+1}(M_r),$$

приходим к (8).

Следствие 2. *Если максимизация в* (8) *выполняется перебором $M_r \in \mathbf{M}$, то при фиксированном r вычислительная сложность алгоритма вычисления ОМП $\widehat{M}_r$, $\widehat{Q} = \widehat{Q}(\widehat{M}_r)$ имеет порядок $O(N^{r+1}s^{r-1})$ и полиномиально зависит от s, N.*

Заметим, что вычислительная сложность идентификации цепи Маркова s-го порядка (с полными связями) имеет порядок $O(N^{s+1})$ и экспоненциально зависит от глубины памяти s.

4. Состоятельность оценок $\widehat{M}_r$ и $\widehat{Q}$

Лемма 2. *Пусть $u_t \in A$ — стационарная цепь Маркова первого порядка с матрицей вероятностей одношаговых переходов $C = (c_{ij})$, $i, j \in A$, $\pi^* = (\pi_i^*)$ — стационарное распределение, а Π^* — $N \times N$ матрица, все строки которой одинаковы и равны π^{*T}, $Z = (z_{ij}) = (\mathbf{I}_N - C + \Pi^*)^{-1}$ — фундаментальная матрица. Пусть далее*

$$n_{ij} = \sum_{t=1}^{n-1} \delta_{u_t,i}\delta_{u_{t+1},j}, \qquad i, j \in A$$

— частота биграммы (i, j) в наблюдаемой последовательности $u_1, \ldots, u_n$ длины n. Тогда при $n \to \infty$ справедливо асимптотическое разложение для ковариаций

$$\operatorname{cov}\{n_{ij}, n_{kl}\} = nc_{ij}(\pi_i^*\delta_{i,k}\delta_{j,l} + c_{kl}(\pi_i^* z_{jk} + \pi_k^* z_{li} - 3\pi_i^*\pi_k^*)) + O(1), \quad i, j, k, l \in A.$$

Доказательство. С помощью эквивалентных преобразований и марковского свойства получаем, что

$$\begin{aligned}\mathbf{E}\{n_{ij}n_{kl}\} &= \sum_{t,t'=1}^{n-1} \mathbf{P}\{u_t = i, u_{t+1} = j, u_{t'} = k, u_{t'+1} = l\} \\ &= (n-1)\delta_{i,k}\delta_{j,l}\pi_i^* c_{ij} + c_{ij}c_{kl}\sum_{s=0}^{n-3}(n-s-2)(\pi_i^*(C^s)_{jk} + \pi_k^*(C^s)_{li}).\end{aligned}$$

Согласно [15], при $m \to \infty$ справедливы матричные разложения

$$\sum_{s=0}^{m} sC^s = \frac{m(m+1)}{2}\Pi^* + O(1)\mathbf{1}_{N\times N},$$

$$\sum_{s=0}^{m} C^s = m\Pi^* + Z + O(\rho^{m+1})\mathbf{1}_{N\times N}.$$

где $0 < \rho < 1$, $\mathbf{L}_{N\times N}$ — $N \times N$ матрица, все элементы которой равны единице.

Учитывая, что $\mathbf{E}\{n_{ij}\} = (n-1)\pi_i^* c_{ij}$, $\operatorname{cov}\{n_{ij}, n_{kl}\} = \mathbf{E}\{n_{ij}n_{kl}\} - \mathbf{E}\{n_{ij}\}\mathbf{E}\{n_{kl}\}$, и выделяя главный член разложения, приходим к доказываемому.

Заметим, что на с. 55 в [14] приведена ошибочная формула для ковариаций, которая дает особенно большие погрешности при малых значениях N.

Примем обозначения (аналогичные обозначениям леммы 2 и теоремы 1) для s-векторной однородной цепи Маркова первого порядка $X_{(t)} = (x_{t-s+1}, \dots, x_t)^T \in A^s$, $t \geqslant s$:

$$P^* = (p^*_{K_s, K'_s}), \quad p^*_{K_s, K'_s} = \delta_{K_s^-, K'_{s-1}} p_{K_s, k'_s}, \quad K_s, K'_s \in A^s,$$

— $N^s \times N^s$ матрица вероятностей одношаговых переходов, где p_{K_s, k'_s} определяются соотношениями (1), $K'_s = (K'_{s-1}, k_s)$, $K_s^- = (k_2, \dots, k_s) \in A^{s-1}$ — усеченный слева мультииндекс; $Z^* = (z^*_{K_s, K'_s})$ — фундаментальная матрица для P^*; $M_{r+1} = (M_r, s+1)$ — шаблон $(r+1)$-го порядка;

$$\begin{aligned}\psi(J_{r+1}, J'_{r+1}; M_r, M'_r) = \sum_{K_{s+1}, K'_{s+1} \in A^{s+1}} &\Pi^*_{K_s} p_{K_{s+1}} p_{K'_{s+1}} (z^*_{K^-_{s+1}, K'_s} - \Pi^*_{K'_s}) \\ &\times (\mathbf{I}\{S_1(K_{s+1}; M_{r+1}) = J_{r+1}, S_1(K'_{s+1}; M'_{r+1}) = J'_{r+1}\} \\ &+ \mathbf{I}\{S_1(K_{s+1}; M'_{r+1}) = J'_{r+1}, S_1(K'_{s+1}; M_{r+1}) = J_{r+1}\}),\end{aligned}$$

$J_{r+1}, J'_{r+1} \in A^{r+1}$, $M_r, M'_r \in \mathbf{M}$. Отметим, что для вычисления матрицы Z^* известны эффективные вычислительные процедуры (см. [15]).

Лемма 3. *Если* ЦМ(s, r) *стационарна, то при* $n \to \infty$ *оценка* $\hat{\mu}_{r+1}$ *состоятельна, т.е.*

$$\hat{\mu}_{r+1}(J_{r+1}; M_r) = \frac{1}{n-s} \nu_{r+1}(J_{r+1}; M_r) \xrightarrow{\mathbf{P}} \mu_{r+1}(J_{r+1}; M_r, M_r^0), \tag{9}$$

где $J_{r+1} \in A^{r+1}$, $M_r \in \mathbf{M}$, *при этом нормированные частоты* $(r+1)$*-грамм*

$$\mu^*_{r+1}(J_{r+1}; M_r) = \frac{1}{\sqrt{n-s}} (\hat{\mu}_{r+1}(J_{r+1}; M_r) - \mu_{r+1}(J_{r+1}; M_r, M_r^0)),$$

где $J_{r+1} \in A^{r+1}$, $M_r \in \mathbf{M}$, *в совокупности распределены асимптотически нормально с нулевыми математическими ожиданиями и ковариациями*

$$\begin{aligned}&\operatorname{cov}\{\mu^*_{r+1}(J_{r+1}; M_r), \mu^*_{r+1}(J'_{r+1}; M'_r)\} \\ &\quad = \sum_{K_{s+1} \in A^{s+1}} \mathbf{I}\{S_1(K_{s+1}; M_{r+1}) = J_{r+1}, S_1(K_{s+1}; M'_{r+1}) = J'_{r+1}\} \Pi^*_{K_s} p_{K_{s+1}} \\ &\qquad + \psi(J_{r+1}, J'_{r+1}; M_r, M'_r) - \mu_{r+1}(J_{r+1}; M_r, M_r^0) \mu_{r+1}(J'_{r+1}; M'_r, M_r^0),\end{aligned} \tag{10}$$

где $J_{r+1}, J'_{r+1} \in A^{r+1}$, $M_r, M'_r \in \mathbf{M}$.

Доказательство. Выразим интересующие нас частоты $(r+1)$-грамм (9) через частоты $(s+1)$-грамм

$$a_{K_{s+1}} = \sum_{t=1}^{n-s} \mathbf{I}\{S_t(X_n; (1, \dots, s+1)) = K_{s+1}\}, \qquad K_{s+1} \in A^{s+1}.$$

Используя (2), получаем, что

$$\nu_{r+1}(J_{r+1}: M_r) = \sum_{K_{s+1}\in A^{s+1}} \mathrm{I}\{S_1(K_{s+1}: M_{r+1}) = J_{r+1}\} a_{K_{s+1}}, \qquad m_{r+1} = s+1, \quad (11)$$

где $J_{r+1} \in A^{r+1}$, $M_r \in \mathbf{M}$. Согласно [14], при $n \to \infty$

$$a_{K_{s+1}}/(n-s) \xrightarrow{\mathbf{P}} \Pi^*_{K_s} p_{K_{s+1}}, \qquad K_{s+1} \in A^{s+1},$$

и нормированные частоты $(s+1)$-грамм

$$\{a^*_{K_{s+1}} = (n-s)^{-1/2}(a_{K_{s+1}} - (n-s)\Pi^*_{K_s} p_{K_{s+1}}): K_{s+1} \in A^{s+1}\}$$

в совокупности распределены асимптотически нормально с нулевыми математическими ожиданиями и ковариациями, вычисляемыми по лемме 2:

$$\begin{aligned}\operatorname{cov}\{a^*_{K_{s+1}}, a^*_{K'_{s+1}}\} &= p_{K_{s+1}} \\ &\times (\Pi^*_{K_s}\delta_{K_{s+1},K'_{s+1}} + p_{K'_{s+1}}(\Pi^*_{K_s} z^*_{K^-_{s+1},K'_s} + \Pi^*_{K'_s} z^*_{K'^-_{s+1},K_s} - 3\Pi^*_{K_s}\Pi^*_{K'_s})).\end{aligned}$$

при $(K_{s+1}, K'_{s+1} \in A^{s+1})$. В силу (5) и линейной связи статистик (11) приходим к (9), (10).

Теорема 4. *Если* ЦМ(s, r) *является стационарной, то при истинном шаблоне* $M_r = M_r^0$ *и* $n \to \infty$ *оценка* (3) *состоятельна:*

$$\hat{Q}(M_r^0) \xrightarrow{\mathbf{P}} Q^0. \quad (12)$$

Доказательство. В силу (11), (5) и свойств частот $\{a_{K_{s+1}}\}$, приведенных в лемме 2,

$$\nu_{r+1}(J_{r+1}: M_r)/(n-s) \xrightarrow{\mathbf{P}} \mu_{r+1}(J_{r+1}: M_r, M_r^0),$$

$$\nu_{r+1}(J_r\cdot: M_r)/(n-s) \xrightarrow{\mathbf{P}} \mu_{r+1}(J_r\cdot: M_r, M_r^0) = \sum_{K_s\in A^s} \mathrm{I}\{S_1(K_s: M_r) = J_r\}\Pi^*_{K_s},$$

где $J_{r+1} \in A^{r+1}$, $n \to \infty$. Так как ЦМ(s, r) стационарна, $\mu_{r+1}(J_r\cdot: M_r^0, M_r^0) > 0$ и $\lim_{n\to\infty}\mathbf{P}\{\nu_{r+1}(J_r\cdot: M_r^0) > 0\} = 1$. Поэтому (12) следует из (3) и теоремы о функциональном преобразовании сходящихся по вероятности случайных последовательностей.

Введем дополнительные обозначения. Пусть $\Phi(z)$ — функция распределения стандартного нормального закона;

$$\beta(J_{r+1}: M_r) = \ln\left(\frac{\mu_{r+1}(J_{r+1}: M_r, M_r^0)}{\mu_{r+1}(J_r\cdot: M_r, M_r^0)\mu_{r+1}(\cdot j_{r+1}: M_r, M_r^0)}\right) - 1,$$

$$\varphi(K_{s+1}, K'_{s+1}: M_r, M'_r) = \beta(S_1(K_{s+1}: M_{r+1}): M_r)\beta(S_1(K'_{s+1}: M'_{r+1}): M'_r):$$

$$\begin{aligned}\Xi = (\Xi_{M_r,M'_r}), \quad \Xi_{M_r,M'_r} = &\sum_{K_{s+1}\in A^{s+1}} \Pi^*_{K_s} p_{K_{s+1}} \varphi(K_{s+1}, K_{s+1}: M_r, M'_r) \\ &- (I_{r+1}(M_r, M_r^0) - 1)(I_{r+1}(M'_r, M_r^0) - 1) \\ &+ \sum_{K_{s+1}, K'_{s+1}\in A^{s+1}} \Pi^*_{K_s} p_{K_{s+1}} p_{K'_{s+1}} (z^*_{K^-_{s+1},K'_s} - \Pi^*_{K'_s}) \\ &\times (\varphi(K_{s+1}, K'_{s+1}: M_r, M'_r) + \varphi(K_{s+1}, K'_{s+1}: M'_r, M_r)),\end{aligned}$$

где $M_r, M'_r \in \mathbf{M}$.

Теорема 5. *Если* ЦМ(s, r) *является стационарной и* $|\Xi| > 0$, *то при* $n \to \infty$ *нормированные уклонения оценок информационного функционала*

$$\{I^*_{r+1}(M_r) = \sqrt{n-s}(\hat{I}_{r+1}(M_r) - I_{r+1}(M_r, M_r^0)) \colon M_r \in \mathbf{M}\}$$

в совокупности распределены асимптотически нормально с нулевыми математическими ожиданиями и ковариационной матрицей Ξ.

Доказательство. По построению, $\hat{I}_{r+1}(M_r)$ — непрерывно дифференцируемое функциональное преобразование асимптотически нормальных частот $(r+1)$-грамм (2):

$$\hat{I}_{r+1}(M_r) = \frac{1}{n-s} \sum_{J_{r+1} \in A^{r+1}} v_{r+1}(J_{r+1} \colon M_r) \times \ln \frac{(n-s) v_{r+1}(J_{r+1} \colon M_r)}{\sum_{J_r \in A^r} v_{r+1}(J_{r+1} \colon M_r) \sum_{j_{r+1} \in A} v_{r+1}(J_{r+1} \colon M_r)}. \tag{13}$$

В силу (13), леммы 3 и условия $|\Xi| \neq 0$ выполнены условия теоремы о функциональном преобразовании асимптотически нормальных случайных последовательностей (см. [16]), что приводит к доказываемому результату.

Лемма 4. *Для любого* $M_r \in \mathbf{M}$

$$I_{r+1}(M_r, M_r^0) \leqslant I_{r+1}(M_r^0, M_r^0).$$

Доказательство. Согласно соотношениям (5), (1) и принятым ранее обозначениям, $\mu_{r+1}(J_{r+1} \colon M_r^0, M_r^0) = \Pi^*(J_r) q_{J_{r+1}}$, $\Pi^*(J_r) = \sum_{K_s \in A^s} \mathbf{I}\{S_1(K_s \colon M_r^0) = J_r\} \Pi^*_{K_s}$ — распределение r-граммы $S_t(X_n \colon M_r^0)$, $J_{r+1} \in A^{r+1}$. Поэтому из (7) следует, что

$$I_{r+1}(M_r^0, M_r^0) = \sum_{J_{r+1} \in A^{r+1}} \Pi^*(J_r) q_{J_{r+1}} \ln(q_{J_{r+1}} / \Pi^*(\cdot, j_{r+1})).$$

Сравним теперь с этой величиной количество информации по Шеннону о будущем символе $x_{t+s} \in A$, содержащейся в s-грамме $(x_t, \ldots, x_{t+s-1}) \in A^s$, включающей r-грамму $S_t(X_n \colon M_r^0)$ как составную часть и вычисляемой аналогично (7):

$$I_{s+1} = \sum_{J_{s+1} \in A^{s+1}} \mu_{s+1}(J_{s+1}) \ln(\mu_{s+1}(J_{s+1})(\mu_{s+1}(J_s \cdot)\mu_{s+1}(\cdot j_{s+1}))^{-1}).$$

Согласно соотношениям (5) и (1) $\mu_{s+1}(J_{s+1}) = \Pi^*_{J_s} q_{S_1(J_s \colon M_r^0), j_{s+1}}$, $\mu_{s+1}(J_s \cdot) = \Pi^*_{J_s}$, $\mu_{s+1}(\cdot j_{s+1}) = \Pi^*(\cdot, j_{s+1})$, поэтому, проводя суммирование по $s - r$ индексам, кроме $S_1(J_s \colon M_r^0)$, заключаем, что справедливо равенство $I_{s+1} = I_{r+1}(M_r^0, M_r^0)$. Любая r-грамма $(x_{t+m_1-1}, \ldots, x_{t+m_r-1}) \in A^r$, порожденная шаблоном $M_r \in \mathbf{M}$, является составной частью (функцией) s-граммы $(x_t, \ldots, x_{t+s-1}) \in A^s$, поэтому по известному свойству количества информации (свойство 6.3.7 из [5]) $I_{r+1}(M_r, M_r^0) \leqslant I_{s+1}$, что и влечет доказываемое неравенство.

Определение 2. Шаблон M_r^0 идентифицируемый, если $I_{r+1}(M_r, M_r^0) = I_{r+1}(M_r^0, M_r^0)$ тогда и только тогда, когда $M_r = M_r^0$.

Теорема 6. *Если* ЦМ(s, r) *является стационарной и шаблон* M_r^0 *идентифицируемый, то при* $n \to \infty$ *оценка* $\widehat{M}_r$, *определяемая* (8), *состоятельна, то есть*

$$\widehat{M}_r \xrightarrow{\mathbf{P}} M_r^0. \tag{14}$$

Доказательство. Из (9), (13) и теоремы о функциональном преобразовании сходящихся по вероятности случайных последовательностей следует сходимость целевых функций в (8), то есть

$$\hat{I}_{r+1}(M_r) \xrightarrow{\mathbf{P}} I_{r+1}(M_r, M_r^0), \qquad M_r \in A^r, \tag{15}$$

а из леммы 4 следует, что $M_r^0 = \arg\max_{M_r} I_{r+1}(M_r, M_r^0)$. Отсюда, из (15) и (8) следует (14).

5. Асимптотическая нормальность $\widehat{Q}$ и проверка гипотез о значении Q^0

В силу теоремы 4 и ограниченности оценок $\hat{q}_{J_{r+1}}(M_r^0) \in [0, 1]$, определяемых (3), из сходимости по вероятности следует состоятельность оценки $\widehat{Q}$ в среднеквадратическом (см. [13]). Поэтому при $n \to \infty$ вариация этой оценки

$$\Delta_n^2 = \mathbf{E}\{\|\widehat{Q} - Q^0\|^2\} \to 0.$$

Теорема 7. *Если* ЦМ(s, r) *является стационарной, то при истинном шаблоне* $M_r = M_r^0$ *случайные величины* $\{q^*_{J_{r+1}} = \sqrt{n-s}(\hat{q}_{J_{r+1}}(M_r^0) - q^0_{J_{r+1}}): J_{r+1} \in A^{r+1}\}$ *при* $n \to \infty$ *в совокупности распределены асимптотически нормально с нулевыми математическими ожиданиями и ковариациями*

$$\operatorname{cov}\{q^*_{J_{r+1}}, q^*_{J'_{r+1}}\} = \mathbf{I}\{J_r = J'_r\} \frac{q^0_{J_{r+1}}(\mathbf{I}\{j_{r+1} = j'_{r+1}\} - q_{J'^0_{r+1}})}{\mu_{r+1}(J_r\cdot : M_r^0, M_r^0)}, \quad J_{r+1}, J'_{r+1} \in A^{r+1}.$$

Доказательство. Так как ЦМ(s, r) стационарна,

$$\lim_{n\to\infty} \mathbf{P}\{\nu_{r+1}(J_r\cdot : M_r^0) > 0\} = 1, \qquad J_r \in A^r,$$

(см. доказательство теоремы 4), поэтому в формуле (3) достаточно рассмотреть случай $\nu_{r+1}(J_r\cdot : M_r^0) > 0$. Таким образом, для оценки $\hat{q}_{J_{r+1}}(M_r^0)$, определяемой (3), справедливы равенства

$$\hat{q}_{J_{r+1}}(M_r^0) = \frac{\hat{\mu}_{r+1}(J_{r+1} : M_r^0)}{\hat{\mu}_{r+1}(J_r\cdot : M_r^0)},$$

$$\frac{\partial \hat{q}_{J_{r+1}}(M_r^0)}{\partial \hat{\mu}_{r+1}(J^{(1)}_{r+1} : M_r^0)} = \mathbf{I}\{J_r = J_r^{(1)}\} \frac{\mathbf{I}\{j_{r+1} = j^{(1)}_{r+1}\} - \hat{q}_{J_{r+1}}(M_r^0)}{\hat{\mu}_{r+1}(J_r\cdot : M_r^0)}.$$

Для нормированных частот $(r+1)$-грамм $\{\mu^*_{r+1}(J_{r+1} : M_r^0): J_{r+1} \in A^{r+1}\}$ согласно лемме 3 асимптотическая ковариация (10) с учетом (1) и эквивалентных преобразований принимает вид

$$\begin{aligned}\operatorname{cov}\{\mu^*_{r+1}(J_{r+1}: M^0_r), \mu^*_{r+1}(J'_{r+1}: M^0_r)\} &= \mathbf{I}\{J_{r+1} = J'_{r+1}\}\mu_{r+1}(J_{r+1}: M^0_r, M^0_r)\\ &\quad - 3\mu_{r+1}(J_{r+1}: M^0_r, M^0_r)\mu_{r+1}(J'_{r+1}: M^0_r, M^0_r)\\ &\quad + q^0_{J_{r+1}} q^0_{J'_{r+1}}(d(J_{r+1}, J'_r) + d(J'_{r+1}, J_r))\\ &= g(J_{r+1}, J'_{r+1}),\\ d(J_{r+1}, J'_r) &= \sum_{K_{s+1}\in A^{s+1}} \sum_{K'_s\in A^s} \Pi^*_{K_s} z^*_{K^-_{s+1}, K'_s} \mathbf{I}\{S_1(K_{s+1}: M^0_{r+1})\\ &= J_{r+1}, S_1(K'_s: M^0_r) = J'_r\},\end{aligned}$$

где $J_{r+1}, J'_{r+1}, \in A^{r+1}$. В силу третьей теоремы непрерывности (см. [17]) случайные величины $\{q^*_{J_{r+1}}: J_{r+1} \in A^{r+1}\}$ в совокупности распределены асимптотически нормально с нулевыми математическими ожиданиями и ковариациями

$$\begin{aligned}\operatorname{cov}\{q^*_{J_{r+1}}, q^*_{J'_{r+1}}\} &= \sum_{J^{(1)}_{r+1}, J^{(2)}_{r+1}\in A^{r+1}} \mathbf{I}\{J_r = J^{(1)}_r\} \frac{\mathbf{I}\{j_{r+1} = j^{(1)}_{r+1}\} - q^0_{J_{r+1}}}{\mu_{r+1}(J_r\cdot : M^0_r, M^0_r)}\\ &\quad \times g(J^{(1)}_{r+1}, J^{(2)}_{r+1})\mathbf{I}\{J'_r = J^{(2)}_r\} \frac{\mathbf{I}\{j'_{r+1} = j^{(2)}_{r+1}\} - q^0_{J'_{r+1}}}{\mu_{r+1}(J'_r\cdot : M^0_r, M^0_r)}\\ &= \sum_{j^{(1)}_{r+1}, j^{(2)}_{r+1}\in A} \frac{(\mathbf{I}\{j_{r+1} = j^{(1)}_{r+1}\} - q^0_{J_{r+1}})(\mathbf{I}\{j'_{r+1} = j^{(2)}_{r+1}\} - q^0_{J'_{r+1}})}{\mu_{r+1}(J_r\cdot : M^0_r, M^0_r)\mu_{r+1}(J'_r\cdot : M^0_r, M^0_r)}\\ &\quad \times g((J_r, j^{(1)}_{r+1}), (J'_r, j^{(2)}_{r+1})).\end{aligned}$$

Учитывая, что $g(\cdot)$ содержит три слагаемых, вычисляем ковариацию:

$$\begin{aligned}&\sum_{j^{(1)}_{r+1}, j^{(2)}_{r+1}\in A} \frac{(\mathbf{I}\{j_{r+1} = j^{(1)}_{r+1}\} - q^0_{J_{r+1}})(\mathbf{I}\{j'_{r+1} = j^{(2)}_{r+1}\} - q^0_{J'_{r+1}})}{\mu_{r+1}(J'_r\cdot : M^0_r, M^0_r)}\\ &\qquad \times \mathbf{I}\{J_r = J'_r, j^{(1)}_{r+1} = j^{(2)}_{r+1}\} q^0_{J_r, j^{(1)}_{r+1}}\\ &\quad = \frac{\mathbf{I}\{J_r = J'_r\}}{\mu_{r+1}(J_r\cdot : M^0_r, M^0_r)} \sum_{j^{(1)}_{r+1}\in A} (\mathbf{I}\{j_{r+1} = j^{(1)}_{r+1}\} - q^0_{J_{r+1}})\\ &\qquad \times (\mathbf{I}\{j'_{r+1} = j^{(1)}_{r+1}\} - q^0_{J'_{r+1}}) q^0_{J_r, j^{(1)}_{r+1}}\\ &\quad = \mathbf{I}\{J_r = J'_r\} q^0_{J_{r+1}}(\mathbf{I}\{j_{r+1} = j'_{r+1}\} - q^0_{J'_{r+1}})(\mu_{r+1}(J_r\cdot : M^0_r, M^0_r))^{-1},\\ &\sum_{j^{(1)}_{r+1}, j^{(2)}_{r+1}\in A} (\mathbf{I}\{j_{r+1} = j^{(1)}_{r+1}\} - q^0_{J_{r+1}})(\mathbf{I}\{j'_{r+1} = j^{(2)}_{r+1}\} - q^0_{J'_{r+1}}) q^0_{J_r, j^{(1)}_{r+1}} q^0_{J'_r, j^{(2)}_{r+1}}\\ &\quad = \sum_{j^{(1)}_{r+1}\in A} (\mathbf{I}\{j_{r+1} = j^{(1)}_{r+1}\} - q^0_{J_{r+1}}) q^0_{J_r, j^{(1)}_{r+1}}\\ &\qquad \times \sum_{j^{(2)}_{r+1}\in A} (\mathbf{I}\{j'_{r+1} = j^{(2)}_{r+1}\} - q^0_{J'_{r+1}}) q^0_{J'_r, j^{(2)}_{r+1}} = 0.\end{aligned}$$

$$\sum_{j_{r+1}^{(1)}, j_{r+1}^{(2)} \in A} \frac{(\mathbf{I}\{j_{r+1} = j_{r+1}^{(1)}\} - q^0_{J_{r+1}})(\mathbf{I}\{j'_{r+1} = j_{r+1}^{(2)}\} - q^0_{J'_{r+1}})}{\mu_{r+1}(J'_r \cdot : M_r^0, M_r^0)\mu_{r+1}(J_r \cdot : M_r^0, M_r^0)}$$

$$\times q^0_{J_r, j_{r+1}^{(1)}} q^0_{J'_r, j_{r+1}^{(2)}} d((J_r, j_{r+1}^{(1)}), J'_r)$$

$$= \sum_{j_{r+1}^{(1)} \in A} \frac{(\mathbf{I}\{j_{r+1} = j_{r+1}^{(1)}\} - q^0_{J_{r+1}})q^0_{J_r, j_{r+1}^{(1)}} d((J_r, j_{r+1}^{(1)}), J'_r)}{\mu_{r+1}(J'_r \cdot : M_r^0, M_r^0)\mu_{r+1}(J_r \cdot : M_r^0, M_r^0)}$$

$$\times \sum_{j_{r+1}^{(2)} \in A} (\mathbf{I}\{j'_{r+1} = j_{r+1}^{(2)}\} - q^0_{J'_{r+1}})q^0_{J'_r, j_{r+1}^{(2)}} = 0.$$

Теорема доказана.

Следствие 3. *В условиях теоремы 7 при $n \to \infty$ имеет место сходимость по распределению, то есть*

$$\mathbf{P}\left\{ \sum_{J_r \in A^r} \sum_{j_{r+1} \in D_{J_r}} \nu_{r+1}(J_r \cdot : M_r^0) \frac{(\hat{q}_{J_{r+1}}(M_r^0) - q^0_{J_{r+1}})^2}{q^0_{J_{r+1}}} \leqslant z \right\} \to G_L(z), \qquad z \geqslant 0,$$

где $G_L(\cdot)$ — функция χ^2-распределения с

$$L = \sum_{J_r \in A^r} d_{J_r}$$

степенями свободы,

$$d_{J_r} = |D_{J_r}| - 1, \qquad D_{J_r} = \{j_{r+1} \in A, : q^0_{J_{r+1}} > 0\}.$$

Доказательство. В силу теоремы 7, ковариационная матрица случайных величин $\{q^*_{J_{r+1}} : J_{r+1} \in A^{r+1}\}$ состоит из N^r блоков размерности $N \times N$, расположенных на диагонали, и нулевых блоков вне диагонали. Поскольку компоненты вектора $q^*(J_r) = (q^*_{J_{r+1}})$, $j_{r+1} \in A$, связаны линейной зависимостью $\sum_{j_{r+1} \in A} q^*_{J_{r+1}} = 0$, диагональные блоки ковариационной матрицы вырождены. Учитывая, что некоторые компоненты Q^0 могут быть равны нулю, получаем, что определитель порядка d_{J_r} каждого из блоков уже отличен от нуля. Таким образом, предельное распределение составного вектора $q^*_L = (q^*_{J_{r+1}})$, $J_r \in A^r$, $j_{r+1} \in \{a_0, a_1, \ldots, a_{d_{J_r}}\}$, $a_0 < a_1 < \ldots < a_{d_{J_r}}$, $a_i \in D_{J_r}$, $i = 0, 1, \ldots, d_{J_r}$, является невырожденным нормальным распределением. А поскольку

$$\nu_{r+1}(J_r \cdot : M_r^0)/(n-s) \xrightarrow{\mathbf{P}} \mu_{r+1}(J_r \cdot : M_r^0, M_r^0),$$

по теореме о квадратичной форме многомерного случайного вектора, имеющего невырожденное нормальное распределение (см. [18]), после некоторых упрощений, получаем требуемый результат.

Лемма 5. *Если* ЦМ(s, r) *является стационарной, то при $n \to \infty$ величина*

$$\mathbf{E}\left\{ \left(\mu_{r+1}(J_r \cdot : M_r^0, M_r^0) \frac{\nu_{r+1}(J_{r+1} : M_r^0)}{n-s} - \mu_{r+1}(J_{r+1} : M_r^0, M_r^0) \frac{\nu_{r+1}(J_r \cdot : M_r^0)}{n-s} \right)^2 \right\}$$

есть $O(1/(n-s))$.

Доказательство. Из (2) следует, что

$$\mathbf{E}\{(\nu_{r+1}(J_{r+1}:M_r^0))^2\} = \sum_{t_1=1}^{n-s}\sum_{t_2=1}^{n-s}\mathbf{P}\{S_{t_1}(X_n:M_{r+1}^0) = J_{r+1},\ S_{t_2}(X_n:M_{r+1}^0) = J_{r+1}\}$$

$$= (n-s)\mu_{r+1}(J_{r+1}:M_r^0,M_r^0) + 2\sum_{t_1=1}^{n-s-1}\sum_{t_2=t_1+1}^{n-s}\mathbf{P}\{S_{t_1}(X_n:M_{r+1}^0) = J_{r+1},\ S_{t_2}(X_n;M_{r+1}^0) = J_{r+1}\}.$$

В силу свойства эргодичности

$$\sum_{K_s\in A^s}\sum_{K_s'\in A^s}\mathbf{P}\{(x_{t_1},\dots,x_{t_1+s}) = (K_s,j_{r+1}),(x_{t_2},\dots,x_{t_2+s}) = (K_s',j_{r+1})\}$$
$$\times\,\mathbf{I}\{S_1(K_s:M_r^0) = J_r\}\mathbf{I}\{S_1(K_s';M_r^0) = J_r\}$$
$$= \sum_{K_s\in A^s}\sum_{K_s'\in A^s}\mathbf{I}\{S_1(K_s;M_r^0) = J_r,\ S_1(K_s';M_r^0) = J_r\}$$
$$\times\,\Pi_{K_s^*}\,p_{K_s,j_{r+1}}(\Pi_{K_s'}^*\,p_{K_s',j_{r+1}} + O(\gamma^{t_2-t_1})),\quad 0<\gamma<1,\quad t_2>t_1,$$

поэтому

$$\mathbf{E}\{(\nu_{r+1}(J_{r+1}:M_r^0))^2\} = (n-s)(n-s-1)(\mu_{r+1}(J_{r+1}:M_r^0,M_r^0))^2 + O(n-s).$$

Аналогичным образом можно показать, что

$$\mathbf{E}\{(\nu_{r+1}(J_r\cdot:M_r^0)/(n-s))^2\} = (\mu_{r+1}(J_r\cdot:M_r^0,M_r^0))^2 + O(1/(n-s)),$$
$$\mathbf{E}\{\nu_{r+1}(J_{r+1}:M_r^0)\,\nu_{r+1}(J_r,\cdot:M_r^0)/(n-s)^2\} = \mu_{r+1}(J_{r+1}:M_r^0,M_r^0)$$
$$\times\,\mu_{r+1}(J_r\cdot:M_r^0,M_r^0) + O(1/(n-s)).$$

Отсюда следует доказываемое асимптотическое соотношение.

Следствие 4. *Если $n\to\infty$, то справедливо асимптотическое разложение вариации*

$$\Delta_n^2 = \frac{1}{n-s}\sum_{J_{r+1}\in A^{r+1}}\frac{(1-q_{J_{r+1}}^0)\,q_{J_{r+1}}^0}{\mu_{r+1}(J_r\cdot:M_r^0,M_r^0)} + o(1/(n-s)). \qquad (16)$$

Доказательство. Воспользуемся представлением

$$(n-s)\Delta_n^2 = \sum_{J_{r+1}\in A^{r+1}}\mathbf{E}\{(\sqrt{n-s}(\hat q_{J_{r+1}}(M_r^0) - q_{J_{r+1}}^0))^2\}.$$

Покажем равномерную ограниченность $\mathbf{E}\{(\sqrt{n-s}(\hat q_{J_{r+1}}(M_r^0)-q_{J_{r+1}}^0))^2\}$, используя лемму 5. Для любых $\varepsilon_{J_{r+1}}\in(0,\mu_{r+1}(J_{r+1}:M_r^0,M_r^0))$, $J_{r+1}\in A^{r+1}$ справедливы соотношения

$$\mathbf{E}\{((\hat q_{J_{r+1}}(M_r^0) - q_{J_{r+1}}^0))^2 \mid |\hat\mu_{r+1}(J_{r+1}:M_r^0) - \mu_{r+1}(J_{r+1}:M_r^0,M_r^0)| \leqslant \varepsilon_{J_{r+1}},\ j_{r+1}\in A\}$$
$$\leqslant (\mu_{r+1}(J_r\cdot:M_r^0,M_r^0)(\mu_{r+1}(J_r\cdot:M_r^0,M_r^0)-\varepsilon))^{-2}O(1/(n-s)) = O(1/(n-s)),$$
$$\varepsilon = \sum_{j_{r+1}\in A}\varepsilon_{J_{r+1}}\in(0,\mu_{r+1}(J_r\cdot:M_r^0,M_r^0)).$$

Таким образом,

$$\mathbf{E}\{((\hat q_{J_{r+1}}(M_r^0) - q^0_{J_{r+1}}))^2\}$$
$$\leqslant \sum_{j_{r+1}\in A} \mathbf{P}\{|\hat\mu_{r+1}(J_{r+1}:M_r^0) - \mu_{r+1}(J_{r+1}:M_r^0, M_r^0)| > \varepsilon_{J_{r+1}}\} + O\left(\frac{1}{n-s}\right).$$

Из леммы 3 следует, что

$$\mathbf{P}\{|\hat\mu_{r+1}(J_{r+1}:M_r^0) - \mu_{r+1}(J_{r+1}:M_r^0, M_r^0)| > \varepsilon_{J_{r+1}}\}$$
$$\approx 2\Phi\left(\frac{-\varepsilon_{J_{r+1}}\sqrt{n-s}}{\sqrt{q^0_{J_{r+1}}(1-q^0_{J_{r+1}})\,\mu_{r+1}(J_r\cdot : M_r^0, M_r^0)}}\right),$$

поэтому

$$(n-s)\mathbf{P}\{|\hat\mu_{r+1}(J_{r+1}:M_r^0) - \mu_{r+1}(J_{r+1}:M_r^0, M_r^0)| > \varepsilon_{J_{r+1}}\} \to 0.$$

Отсюда следует, что

$$\mathbf{E}\{(\sqrt{n-s}(\hat q_{J_{r+1}}(M_r^0) - q^0_{J_{r+1}}))^2\} = O(1).$$

Заметим также, что

$$\mathbf{E}\{(\sqrt{n-s}(\hat q_{J_{r+1}}(M_r^0) - q^0_{J_{r+1}}))^k \mid |\sqrt{n-s}(\hat q_{J_{r+1}}(M_r^0) - q^0_{J_{r+1}})| > \sqrt{n-s}\} \equiv 0$$

для любого $n > s$ и $k \in \mathbf{N}$. Воспользовавшись теоремой 7 и теоремой о сходимости интегралов от равномерно интегрируемых функций (см. [13]), получаем (16).

Используя следствие 3, построим критерий проверки гипотез $H_0\colon Q^0 = Q_0$ против альтернативы общего вида $H_1 = \bar H_0$, где $Q_0 = (q_{0J_{r+1}})$ — некоторая фиксированная стохастическая матрица. Решающее правило для проверки H_0 против альтернативы H_1 основано на статистиках (3) и состоит в том, что

$$\text{принимается гипотеза } \begin{cases} H_0, & \text{если } \rho \leqslant \Delta, \\ H_1, & \text{если } \rho > \Delta, \end{cases} \tag{17}$$

где

$$\rho = \sum_{J_r\in A^r} \sum_{j_{r+1}\in D_{J_r}} \nu_{r+1}(J_r\cdot : M_r^0)\frac{(\hat q_{J_{r+1}}(M_r^0) - q_{0J_{r+1}})^2}{q_{0J_{r+1}}},$$

$\Delta > 0$ — некоторый порог.

Следствие 5. *Пусть $\varepsilon \in (0, 1)$ и $\Delta = G_L^{-1}(1-\varepsilon)$ — квантиль уровня $1-\varepsilon$ стандартного χ^2-распределения с L степенями свободы. Тогда предельный при $n \to \infty$ размер критерия* (17) *равен ε.*

Доказательство. В силу следствия 3 вероятность ошибки первого рода

$$\alpha = \mathbf{P}\{\rho > \Delta \mid H_0\} = 1 - \mathbf{P}\{\rho \leqslant \Delta \mid H_0\} \to 1 - G_L(\Delta) = \varepsilon.$$

Приведем еще эквивалентный вид критерия (17):

$$\text{принимается гипотеза} \begin{cases} H_0, & \text{если } P \geqslant \varepsilon, \\ H_1, & \text{если } P < \varepsilon, \end{cases}$$

где $P = 1 - G_L(\rho)$ — так называемое P-значение χ^2-распределения с L степенями свободы.

Следствие 6. *Если*

$$H_{1n}: q_{1J_{r+1}} = q_{0J_{r+1}} + \frac{c_{J_{r+1}}}{\sqrt{n-s}} + O(1/n), \qquad n \to \infty,$$

— семейство контигуальных альтернатив, где константы $\{c_{J_{r+1}}: J_{r+1} \in A^{r+1}\}$ *таковы, что*

$$\sum_{j_{r+1} \in A} c_{J_{r+1}} = 0, \qquad \sum_{J_r \in A^r} \sum_{j_{r+1} \in D_{J_r}} c^2_{J_{r+1}} > 0,$$

то при $\Delta = G_L^{-1}(1-\varepsilon)$ *асимптотическая мощность теста* (17)

$$w \to 1 - G_{L,a}(G_L^{-1}(1-\varepsilon)), \tag{18}$$

где $G_{L,a}(\cdot)$ *— функция нецентрального распределения* χ^2 *с* L *степенями свободы и параметром нецентральности*

$$a = \sum_{J_r \in A^r} \sum_{j_{r+1} \in D_{J_r}} \mu_{r+1}(J_r\cdot; M_r^0; M_r^0) c^2_{J_{r+1}}/q_{0J_{r+1}} > 0.$$

Доказательство. Поскольку

$$\frac{\nu_{r+1}(J_r\cdot; M_r^0)}{n-s} \xrightarrow{\mathbf{P}} \mu_{r+1}(J_r\cdot; M_r^0, M_r^0),$$

в силу теоремы 7 в случае справедливости альтернативы H_{1n} при $n \to \infty$ случайная величина ρ имеет в пределе нецентральное распределение χ^2 (см. способы его вычисления в [18, 19]) с L степенями свободы и параметром нецентральности

$$\begin{aligned} &\lim_{n\to\infty} \sum_{J_r \in A^s} \sum_{j_{r+1} \in D_{J_r}} \left(\mathbf{E}\left\{(\hat{q}_{J_{r+1}} - q_{0J_{r+1}})\sqrt{(n-s)\mu_{r+1}(J_r\cdot; M_r^0, M_r^0)/q_{0J_{r+1}}}\right\}\right)^2 \\ &= \lim_{n\to\infty} (n-s) \sum_{J_r \in A^s} \sum_{j_{r+1} \in D_{J_r}} \frac{\mu_{r+1}(J_r\cdot; M_r^0, M_r^0)}{q_{0J_{r+1}}} (\mathbf{E}\{\hat{q}_{J_{r+1}}\} - q_{0J_{r+1}})^2 \\ &= \sum_{J_r \in A^r} \sum_{j_{r+1} \in D_{J_r}} \mu_{r+1}(J_r\cdot; M_r^0, M_r^0) c^2_{J_{r+1}}/q_{0J_{r+1}} = a. \end{aligned}$$

Таким образом, мощность теста

$$w = \mathbf{P}\{1 - G_L(\rho) < \varepsilon \mid H_1\} = \mathbf{P}\{\rho > G_L^{-1}(1-\varepsilon) \mid H_1\} \to 1 - G_{L,a}(G_L^{-1}(1-\varepsilon)),$$

что совпадает с (18).

6. Алгоритмы вычисления оценки шаблона $\widehat{M}_r$

Для вычисления оценки $\widehat{M}_r$ шаблона M_r^0 согласно (8) при известном истинном значении r^0 числа связей r предлагаются два алгоритма: алгоритм A1 полного перебора значений целевой функции (8) и алгоритм A2 наращивания шаблона, обеспечивающий сокращение перебора. В случае, когда r^0 известно с точностью до числового промежутка: $r_- \leqslant r^0 \leqslant r_+$, $1 \leqslant r_- < r_+ \leqslant s$, для совместного оценивания r^0, M_r^0 предлагается алгоритм A3 сокращения шаблона.

При описании алгоритмов A2, A3 мы используем вспомогательные обозначения. Для некоторого шаблона r-го порядка $M_r = (m_1, m_2, \dots, m_r) \in \mathbf{M}$, $2 \leqslant r < s$, обозначим через $F^+(M_r)$ подмножество $s - r$ всевозможных шаблонов $(r+1)$-го порядка, получающихся вставкой в строку $m_1, m_2, \dots, m_r$ одного из новых символов $m \in \{1, 2, \dots, s\} \setminus \{m_1, m_2, \dots, m_r\}$; символ вставляется так, чтобы новая строка символов $m'_1, m'_2, \dots, m'_{r+1}$ обладала свойством шаблона: $1 = m'_1 < \dots < m'_{r+1} \leqslant s$. Аналогично, через $F^-(M_r)$ обозначим подмножество $r - 1$ всевозможных шаблонов $(r-1)$-го порядка, которые получаются вычеркиванием одного из символов $m \in \{m_2, m_3, \dots, m_r\}$ за исключением первого.

Алгоритм A2 наращивания шаблона заключается в последовательном вычислении шаблонов $\widehat{M}_{r_-}, \widehat{M}_{r_-+1}, \dots, \widehat{M}_{r^0}$, где $r_- \in \{1, 2, \dots, r^0\}$ — наименьший размер шаблона, задаваемый исходя из имеющихся вычислительных ресурсов. Вначале при $r = r_-$ алгоритмом A1 вычисляется начальный шаблон $\widehat{M}_{r_-}$ порядка r_-; затем осуществляется наращивание этого шаблона по рекуррентной формуле

$$\widehat{M}_{r_-+i} = \arg \max_{M_{r_-+i} \in F^+(\widehat{M}_{r_-+i-1})} \hat{I}_{r_-+i+1}(M_{r_-+i}), \qquad i = 1, 2, \dots, r^0 - r_-.$$

Наибольшее быстродействие алгоритма A2, очевидно, достигается при $r_- = 1$. Этот подход аналогичен алгоритму расширения пространства признаков в задачах распознавания образов, который как известно, может приводить к потере истинной гипотезы.

Алгоритм A3 базируется на очевидном свойстве вложенности моделей цепей Маркова:

$$\text{ЦМ}(s, 1) \subset \text{ЦМ}(s, 2) \subset \dots \subset \text{ЦМ}(s, s),$$

в силу которого шаблон связей $M_{r^0+1}^0$ содержит $M_{r^0}^0$. Вначале при $r = r_+$ алгоритмом A1 вычисляется начальный шаблон $\widehat{M}_{r_+}$ порядка r_+; затем осуществляется сокращение этого шаблона по рекуррентной формуле

$$\widehat{M}_{r_+-i} = \arg \max_{M_{r_+-i} \in F^-(\widehat{M}_{r_+-i+1})} \hat{I}_{r_+-i+1}(M_{r_+-i}), \qquad i = 1, 2, \dots, r_+ - r_-.$$

Затем на основе $\{\widehat{M}_r, \ \hat{I}_{r+1}(\widehat{M}_r) : r \in \{r_-, r_- + 1, \dots, r_+\}\}$ строится искомая оценка $\widehat{M}_{\hat{r}}$, где

$$\hat{r} = \arg \max_{r \in \{r_-, r_-+1, \dots, r_+-1} \}(\hat{I}_{r+1}(\widehat{M}_r) - \hat{I}_r(\widehat{M}_{r-1})), \hat{I}_1(\widehat{M}_0) = 0.$$

Подобный способ используется в прикладной статистике в пошаговой регрессии для определения информативного множества предикторов [20].

7. Вероятность ошибки распознавания шаблона

Будем исследовать статистическую оценку шаблона $\widehat{M}_r$, вычисленную алгоритмом A1 полного перебора, в предположении идентифицируемости шаблона M_r^0. Множество шаблонов $\mathbf{M}$ мощности $K = |\mathbf{M}|$ разобьем на K^*, $2 \leqslant K^* \leqslant K$, классов эквивалентности:

$$\mathbf{M} = \bigcup_{k=0}^{K^*-1} \mathbf{M}^{(k)}, \qquad M^{(k)} \cap M^{(l)} = \varnothing, \qquad k \neq l,$$

$$M^{(k)} = \{M_r^{(k,1)}, \ldots, M_r^{(k,K_k^*)}\}, \qquad 1 \leqslant K_k^* < K, \quad \sum_{k=0}^{K^*-1} K_k^* = K,$$

так, что в каждом классе информационный функционал (7) для всех шаблонов одинаков и выполняются неравенства

$$I_{r+1}(M_r^{(k,1)}: M_r^0) > I_{r+1}(M_r^{(k+1,1)}: M_r^0), \qquad k = 0, 1, \ldots, K^* - 2.$$

При этом в силу свойства идентифицируемости истинного шаблона M_r^0 и леммы 4 нулевой класс $\mathbf{M}^{(0)}$ содержит лишь один шаблон M_r^0: $K_0^* = 1$. Пусть

$$\xi^{(k,j_k)} = \hat{I}_{r+1}(M_r^{(k,j_k)}), \quad j_k \in \{1, 2, \ldots, K_k^*\}, \quad k \in \{0, 1, \ldots, K^* - 1\},$$

тогда вероятность ошибки при распознавании шаблона с помощью решающего правила (8) равна

$$\begin{aligned} P_0 &= \mathbf{P}\{\widehat{M}_r \neq M_r^0\} = \mathbf{P}\{\widehat{M}_r \neq M_r^{(0,1)}\} \\ &= \mathbf{P}\left\{\hat{I}_{r+1}(M_r^{(0,1)}) < \max_{1 \leqslant j_k \leqslant K_k^*,\, 1 \leqslant k \leqslant K^*-1} \hat{I}_{r+1}(M_r^{(k,j_k)})\right\} \\ &= \mathbf{P}\left\{\xi^0 < \max_{1 \leqslant j_k \leqslant K_k^*,\, 1 \leqslant k \leqslant K^*-1} \xi^{(k,j_k)}\right\}, \end{aligned}$$

где $\xi^0 = \xi^{(0,1)}$. Здесь полагается, что в случае неединственности оценки (8) ошибка распознавания отсутствует, если M_r^0 входит в множество эквивалентных оценок шаблона $\widehat{M}_r$.

Для вероятности ошибки при различении пары шаблонов $M_r^{(0,1)}$, $M_r^{(k,j_k)}$ для некоторых фиксированных $j_k \in \{1, 2, \ldots, K_k^*\}$ и $k \in \{1, 2, \ldots, K^* - 1\}$ введем обозначение

$$\begin{aligned} P_{k,j_k} &= \mathbf{P}\{\hat{I}_{r+1}(M_r^{(0,1)}) < I_{r+1}(M_r^{(k,j_k)})\} \\ &= \mathbf{P}\{\xi^0 < \xi^{(k,j_k)}\} = 1 - \mathbf{P}\{\xi^0 \geqslant \xi^{(k,j_k)}\} = 1 - \mathbf{P}\{\xi^{(k,j_k)} - \xi^0 \leqslant 0\}. \end{aligned}$$

Теорема 8. *Если используется алгоритм (8) распознавания шаблона, то для вероятности ошибки справедливо двухстороннее неравенство*

$$\begin{aligned} &P_- \leqslant P_0 \leqslant P_+, \\ &P_- = \max_{1 \leqslant j_k \leqslant K_k^*,\, 1 \leqslant k \leqslant K^*-1} P_{k,j_k}, \\ &P_+ = \min\left\{\sum_{k=1}^{K^*-1} \sum_{j_k=1}^{K_k^*} P_{k,j_k},\ 1\right\}. \end{aligned} \qquad (19$$

Доказательство. С одной стороны, по свойствам вероятности

$$1 - P_0 = \mathbf{P}\left(\bigcap_{1 \leqslant j_k \leqslant K_k^*,\ 1 \leqslant k \leqslant K^*-1} \{\xi^{(k,j_k)} - \xi^0 \leqslant 0\}\right)$$
$$\leqslant \min_{1 \leqslant j_k \leqslant K_k^*,\ 1 \leqslant k \leqslant K^*-1} \mathbf{P}\{\xi^{(k,j_k)} - \xi^0 \leqslant 0\} = 1 - \max_{1 \leqslant j_k \leqslant K_k^*,\ 1 \leqslant k \leqslant K^*-1} P_{k,j_k},$$

откуда следует оценка снизу $P_0 \geqslant P_-$.

С другой стороны, по свойствам вероятности

$$P_0' = \mathbf{P}\left(\bigcup_{1 \leqslant j_k \leqslant K_k^*,\ 1 \leqslant k \leqslant K^*-1} \{\xi^{(k,j_k)} - \xi^0 > 0\}\right)$$
$$\leqslant \sum_{k=1}^{K^*-1} \sum_{j_k=1}^{K_k^*} \mathbf{P}\{\xi^{(k,j_k)} - \xi^0 > 0\} = \sum_{k=1}^{K^*-1} \sum_{j_k=1}^{K_k^*} P_{k,j_k} = P_+.$$

Объединяя эти два неравенства, получаем (19).

Для приближенного вычисления вероятностей $\{P_{k,j_k}\}$, входящих в (19), воспользуемся асимптотической нормальностью $\{\xi^{(k,j_k)}\}$, вытекающей из теоремы 5, и получим, что

$$P_{k,j_k} = 1 - \mathbf{P}\{\xi^{(k,j_k)} - \xi^0 \leqslant 0\} \approx \Phi(\sqrt{n-s}\, e_{k,j_k}/\sigma_{k,j_k}), \tag{20}$$

где

$$\sigma_{k,j_k}^2 = \Xi_{M_r^0, M_r^0} - 2\Xi_{M_r^0, M_r^{(k,j_k)}} + \Xi_{M_r^{(k,j_k)}, M_r^{(k,j_k)}} > 0,$$
$$e_{k,j_k} = I_{r+1}(M_r^{(k,j_k)}; M_r^0) - I_{r+1}(M_r^0; M_r^0) < 0.$$

Относительно точности приближения (20) справедливы замечания, приведенные в разделе 4.5 в [13].

Рассмотрим еще один подход к приближенному вычислению P_0. Определим ковариационную матрицу (K^*-1)-го порядка $\Xi^* = (\Xi_{kl}^*)$, где

$$\Xi_{kl}^* = \Xi_{M_r^0, M_r^0} + \Xi_{M_r^{(k,1)}, M_r^{(l,1)}} - \Xi_{M_r^0, M_r^{(k,1)}} - \Xi_{M_r^0, M_r^{(l,1)}},$$

$k, l = 1, 2, \ldots, K^* - 1$, и функцию распределения (K^*-1)-мерного нормального закона с нулевым вектором математического ожидания и ковариационной матрицей Ξ^*: $\Phi_{K^*-1}(z_1, \ldots, z_{K^*-1}; \Xi^*)$, $z_k \in \mathbf{R}$, $k = 1, \ldots, K^*-1$. При этом $\Xi_{M_r^{(k,1)}, M_r^{(l,1)}}$, $k, l = 1, 2, \ldots, K^*-1$, определены в теореме 5, так что Ξ^* является главной подматрицей матрицы Ξ.

Лемма 6. *Пусть $\zeta_n^{(1)}, \zeta_n^{(2)}, \zeta_n^{(3)}$, $n \in \mathbf{N}$, — случайные последовательности такие, что $\zeta_n^{(2)}$ и $\zeta_n^{(3)}$ эквивалентны в смысле сходимости по вероятности:*

$$\zeta_n^{(3)} - \zeta_n^{(2)} \xrightarrow{\mathbf{P}} 0,$$

и имеет место сходимость по распределению:

$$\mathbf{P}\{\zeta_n^{(1)} \leqslant z_1,\ \zeta_n^{(2)} \leqslant z_2\} \to F(z_1, z_2), \qquad z_1, z_2 \in \mathbf{R}.$$

причем $F(\cdot)$ — непрерывная функция. Тогда при $n \to \infty$

$$\mathbf{P}\{\zeta_n^{(1)} \leqslant z_1, \zeta_n^{(2)} \leqslant z_2, \zeta_n^{(3)} \leqslant z_2\} \to F(z_1, z_2), \qquad z_1, z_2 \in \mathbf{R}.$$

Доказательство. По формуле полной вероятности

$$\mathbf{P}\{\zeta_n^{(1)} \leqslant z_1, \zeta_n^{(2)} \leqslant z_2, \zeta_n^{(3)} \leqslant z_2\} = \mathbf{P}\{\zeta_n^{(1)} \leqslant z_1, \zeta_n^{(2)} \leqslant z_2\} - \gamma_n,$$

где

$$\begin{aligned}\gamma_n &= \mathbf{P}\{\zeta_n^{(1)} \leqslant z_1, \zeta_n^{(2)} \leqslant z_2, \zeta_n^{(3)} > z_2\} \\ &= \mathbf{P}\{\zeta_n^{(1)} \leqslant z_1, \zeta_n^{(2)} \leqslant z_2, \zeta_n^{(3)} > z_2, |\zeta_n^{(1)} - \zeta_n^{(2)}| \leqslant \varepsilon\} \\ &\quad + \mathbf{P}\{\zeta_n^{(1)} \leqslant z_1, \zeta_n^{(2)} \leqslant z_2, \zeta_n^{(3)} > z_2, |\zeta_n^{(3)} - \zeta_n^{(2)}| > \varepsilon\} \\ &\leqslant \mathbf{P}\{\zeta_n^{(1)} \leqslant z_1, z_2 - \varepsilon < \zeta_n^{(2)} \leqslant z_2\} + \mathbf{P}\{|\zeta_n^{(3)} - \zeta_n^{(2)}| > \varepsilon\} \\ &\to F(z_1, z_2) - F(z_1, z_2 - \varepsilon),\end{aligned}$$

для любого $\varepsilon > 0$. Отсюда в силу произвольности ε и непрерывности $F(\cdot)$ получаем, что $\gamma_n \to 0$, это и проводит к доказываемому результату.

Теорема 9. *Если* ЦМ(s, r) *является стационарной и матрица Ξ^* невырождена, то при $n \to \infty$ для вероятности ошибки распознавания шаблона справедливо приближение*

$$P_0 \approx 1 - \Phi_{K^*-1}(\sqrt{n-s}(I_{r+1}(M_r^0) - I_{r+1}(M_r^{(1,1)})), \\ \ldots, \sqrt{n-s}(I_{r+1}(M_r^0) - I_{r+1}(M_r^{(K^*-1,1)})); \Xi^*). \quad (21)$$

Доказательство. В принятых обозначениях

$$1 - P_0 = \mathbf{P}\left(\bigcap_{1 \leqslant j_k \leqslant K_k^*,\ 1 \leqslant k \leqslant K^*-1} \{\xi^{(k, j_k)} - \xi^0 \leqslant 0\}\right).$$

Поэтому в силу леммы 6 при вычислении указанной вероятности достаточно в k-м классе эквивалентности, $1 \leqslant k \leqslant K^* - 1$, выбрать лишь одного представителя $\xi^{(k,1)}$. В силу теоремы 5 случайные величины $\{\xi^{(k,1)}: k \in \{0, 1, \ldots, K^* - 1\}\}$ распределены асимптотически нормально, что влечет (21). Заметим, что невырожденность матрицы Ξ^* вытекает из невырожденности матрицы Ξ.

8. Численные результаты

Прежде всего, остановимся на вычислительной сложности алгоритмов A1, A2, A3. Проведенный на `Athlon XP 1600+` анализ быстродействия алгоритмов вычисления шаблона $\widehat{M}_r$ дал следующие оценки затрат машинного времени $t^{(1)}, t^{(2)}, t^{(3)}$ в секундах для A1, A2, A3 соответственно:

$$t^{(1)} \approx 1.19 \cdot 10^{-7} \binom{s-1}{r-1} (N^{r+1} + nr) = O(N^{r+1} s^{r-1} + ns^r);$$

Таблица 1.

№	N	s	r	K^*	n	A1 P_0 способ 1	A1 P_0 способ 2	A1 $\widehat{P}_0$	A2 $\widehat{P}_0$	A3 $\widehat{P}_0$
1	2	4	2	2	5000	[0, 0,9831]	0,2642	0,4813	0,4813	0,4904
					20000	[0, 0,9662]	0,1036	0,1661	0,1661	0,1817
					50000	[0, 0,9467]	0,0230	0,0181	0,0181	0,0210
2	2	4	2	2	3000	[0, 0,4168]	0,0026	0,0004	0,0004	0,0005
3	2	8	2	5	10000	[0, 0,1968]	0	0	0	0
4	2	8	2	4	10000	[0, 1]	0,0470	0,0491	0,0491	0,0525
5	2	6	4	10	3000	[0, 0,0053]	0	0	0	0
6	3	4	3	3	5000	[0, 0]	0	0	0	0
7	4	4	2	3	8000	[0, 0]	0	0	0	0

$$t^{(2)} \approx 3.01 \cdot 10^{-7} \left(\binom{s-1}{r_- - 1} (N^{r_-+1} + nr_-) + \sum_{i=r_-+1}^{r} (s-i+1)(N^{i+1} + ni) \right)$$
$$= O(N^{r_-+1} s^{r_- - 1} + ns^{r_-} + N^{r+2} s + nsr^2);$$
$$t^{(3)} \approx 1.04 \cdot 10^{-7} \left(\binom{s-1}{r_+ - 1} (N^{r_++1} + nr_+) + \sum_{i=r_-}^{r_+-1} i(N^{i+1} + ni) \right)$$
$$= O(N^{r_++1} s^{r_+-1} + ns^{r_+} + N^2(N^{r_-} + r_+ N^{r_+} + nr_-^2 + nr_+^3)).$$

Формульные множители в этих выражениях найдены прямым подсчетом числа основных компьютерных операций каждого алгоритма, а числовые множители находились методом наименьших квадратов на основе выборки экспериментально зарегистрированных затрат машинного времени алгоритмов при различных сочетаниях N, s, r, n, r_-, r_+. По преимуществу в быстродействии алгоритмы упорядочены следующим образом: A2, A1, A3. Так, например, при $N = 2$, $r = 3$, $s = 128$, $n = 5\,000$, $t^{(1)} \approx 13.93$, $t^{(2)} \approx 0.93$ ($r_- = 1$), $t^{(3)} \approx 676.78$ ($r_- = 3$, $r_+ = 4$), что не противоречит полученным экспериментальным значениям затрат машинного времени: $\hat{t}^{(1)} = 13.70$, $\hat{t}^{(2)} = 1.01$, $\hat{t}^{(3)} = 616.10$.

По методу Монте–Карло при числе прогонов 10^4 для алгоритмов A1, A2, A3 получены значения частоты ошибки $\widehat{P}_0$ распознавания шаблона. Для алгоритма A1 также получены асимптотические приближения для вероятности ошибки P_0 с помощью двух выше изложенных способов. Для фиксированных значений N, s и r проведено несколько серий экспериментов с различными матрицами Q^0, элементы которой генерировались с помощью генератора равномерно распределенных на $[0, 1]$ псевдослучайных чисел с последующей нормировкой каждой строки. Результаты семи серий экспериментов представлены в табл. 1, где способ 1 — это интервальная оценка (19) с приближением (20), способ 2 — оценка (21).

При $r = 2$ различий между алгоритмом A1 полного перебора и алгоритмом A2 наращивания шаблона нет, поэтому вероятности ошибок совпадают. Достаточно большая

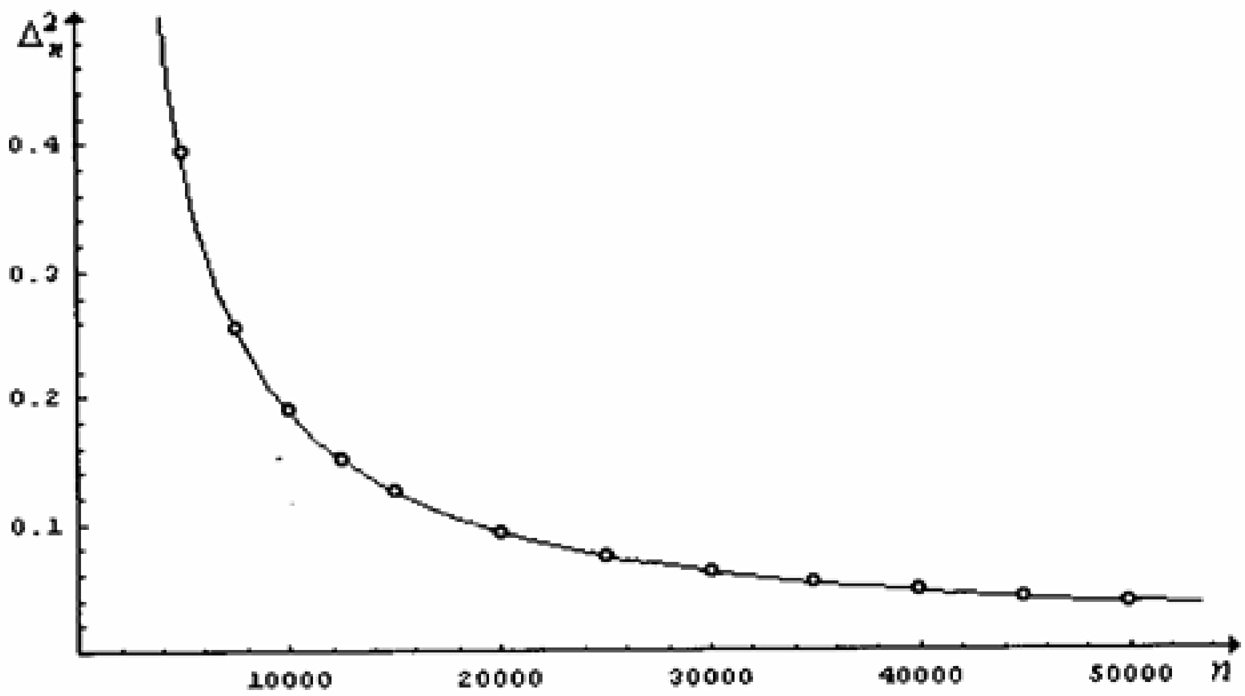

Рис. 1. Зависимость Δ_n^2 от n при $N = 2$, $s = 256$, $r = 6$

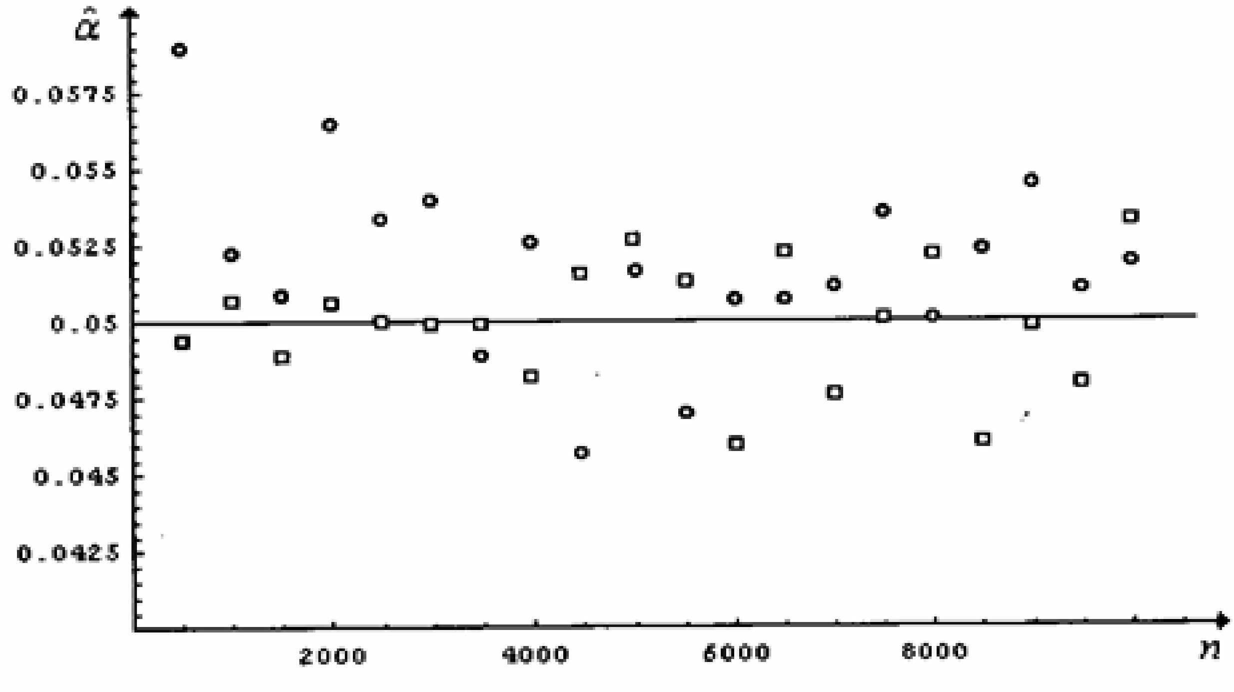

Рис. 2. Экспериментальные значения вероятности ошибки первого рода при $N = 2$, $s = 256$, $r = 4$, $\varepsilon = 0{,}05$ для двух различных матриц Q^0

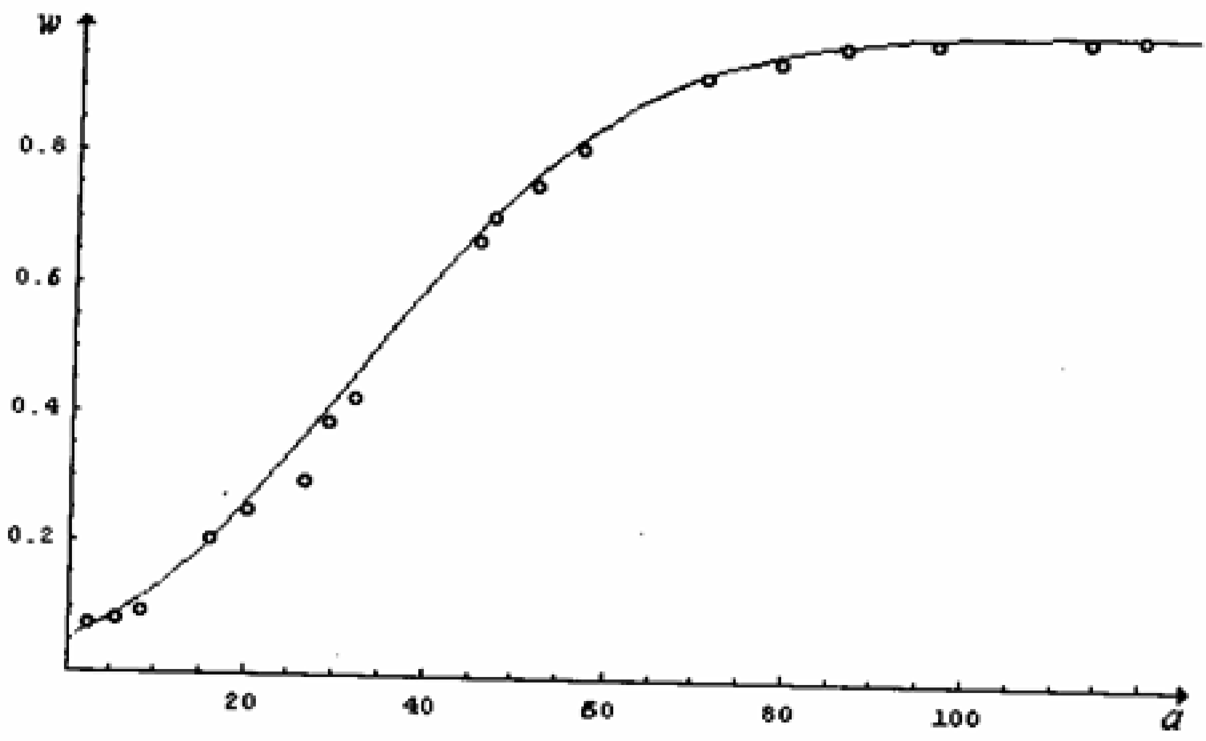

Рис. 3. Зависимость w от a при $N = 4$, $s = 256$, $r = 3$, $n = 10\,000$, $\varepsilon = 0.05$

вероятность ошибки в серии №1 объясняется малой длительностью наблюдения n и специфическим видом матрицы Q^0 (использовался шаблон $M_2^0 = (1, 4)$):

$$Q^0 = \begin{pmatrix} 0.4627 & 0.5373 \\ 0.4459 & 0.5541 \\ 0.7327 & 0.2673 \\ 0.7153 & 0.2847 \end{pmatrix}.$$

Поскольку

$$\begin{aligned} \mathbf{P}\{x_{s+1} = 0 \mid x_{m_1^0} = 0, x_{m_2^0} = 0\} &\approx \mathbf{P}\{x_{s+1} = 0 \mid x_{m_1^0} = 0, x_{m_2^0} = 1\} \\ &\approx \mathbf{P}\{x_{s+1} = 0 \mid x_{m_1^0} = 0\} = 0.45, \\ \mathbf{P}\{x_{s+1} = 0 \mid x_{m_1^0} = 1, x_{m_2^0} = 0\} &\approx \mathbf{P}\{x_{s+1} = 0 \mid x_{m_1^0} = 1, x_{m_2^0} = 1\} \\ &\approx \mathbf{P}\{x_{s+1} = 0 \mid x_{m_1^0} = 1\} = 0.72, \end{aligned}$$

параметр r в этом примере фактически равняется 1, а не 2. Видно, что при увеличении длительности наблюдений n истинный шаблон распознается с наперед заданной точностью.

Численные результаты показывают, что наилучшим по точности является алгоритм A1 полного перебора, затем алгоритм A3 сокращения шаблона и алгоритм A2 наращивания шаблона.

В численных экспериментах исследовалась также вариация оценки $\hat{Q}$:

$$\Delta_n^2 = \mathbf{E}\{\|\hat{Q} - Q^0\|^2\}.$$

На рис. 1 для ЦМ(256, 6) приведен график зависимости Δ_n^2 от n (кривая вычислена теоретически с помощью главного члена разложения (16), кружки — экспериментальные значения).

На рис. 2 для ЦМ(256, 4) представлены экспериментальные значения вероятности ошибки первого рода $\hat{\alpha}$ критерия (17) для различных значений длительности наблюдения n и прямая $\alpha = \varepsilon = 0.05$, соответствующая заданному уровню значимости.

На рис. 3 для ЦМ(256, 3) изображена теоретическая кривая зависимости мощности критерия (17) от параметра нецентральности a и соответствующие экспериментальные значения.

Авторы глубоко благодарны А. М. Зубкову за уточнения постановок задач и критические замечания, способствовавшие улучшению статьи.

Список литературы

1. Fan J., Yao Q., *Nonlinear time series.* Springer, Berlin, 2003.
2. Collett D., *Modelling binary data.* Chapman & Hall, Boca Raton, 2003.
3. Yaffee R. A., McGee M., *Time series analysis and forecasting.* Academic Press, San Diego, 2000.
4. Waterman M. S., *Mathematical methods for DNA sequences.* CRC Press, Boca Raton, 1989.
5. Харин Ю. С., Берник В. И., Матвеев Г. В., Агиевич С. В., *Математические и компьютерные основы криптологии.* Новое знание, Минск, 2003.
6. Зубков А. М., Датчики псевдослучайных чисел и их применения. В сб.: *Труды II Международной научной конференции «Математика безопасности информационных технологий».* МГУ, Москва, 2003, с. 200–206.
7. Дуб Д., *Вероятностные процессы.* ИЛ, Москва, 1965.
8. Raftery A. E., A model for high-order Markov chains. *J. Royal Statist. Soc.* (1985) **47**, №3, 528–539.
9. Харин Ю. С. Цепи Маркова с r-частичными связями и их статистическое оценивание. *Докл. НАН Беларуси* (2004) **48**, №1, 40–44.
10. Тихомирова М. И., Чистяков В. П., О двух статистиках типа хи-квадрат, построенных по частотам цепочек состояний сложной цепи Маркова. *Дискретная математика.* (2003) **15**, №2. 149–159.
11. Максимов Ю. И., О цепях Маркова, связанных с двоичными регистрами сдвига со случайными элементами. *Труды по дискретной математике* (1997) **1**, 203–220.
12. Лидл Р., Нидеррайтер Г. *Конечные поля.* Мир, Москва, 1988.
13. Королюк В. С., Портенко Н. И., Скороход А. В., Турбин А. Ф. *Справочник по теории вероятностей и математической статистике.* Наука, Москва, 1985.
14. Basawa, I. V., Prakasa Rao B. L. S., *Statistical inference for stochastic processes.* Academic Press London, 1980.
15. Кемени Д., Снелл Д. *Конечные цепи Маркова.* Физматлит, Москва, 1970.
16. Serfling R. J., *Approximation theorems of mathematical statistics.* Wiley, New York, 1980.
17. Боровков А. А., *Математическая статистика.* Наука, Москва, 1984.
18. Ивченко Г. И., Медведев Ю. И., *Математическая статистика.* Высшая школа, Москва, 1984
19. Большев Л. Н., Смирнов Н. В., *Таблицы математической статистики.* Наука, Москва, 1983.
20. Айвазян С. А., Енюков И. С., Мешалкин Л. Д., *Прикладная статистика. Исследование зависимостей.* Финансы и статистика, Москва, 1985.

Статья поступила 11.04.2006