ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)
УДК 519.217.2;519.233.22
https://doi.org/10.29235/1561-2430-2022-58-3-280-291

Поступила в редакцию 24.06.2022
Received 24.06.2022

В. А. Волошко, Ю. С. Харин

Научно-исследовательский институт прикладных проблем математики и информатики, Белорусский государственный университет, Минск, Республика Беларусь

ДИСКРЕТНЫЕ ВРЕМЕННЫЕ РЯДЫ НА ОСНОВЕ ЭКСПОНЕНЦИАЛЬНОГО СЕМЕЙСТВА С МНОГОМЕРНЫМ ПАРАМЕТРОМ И ИХ ВЕРОЯТНОСТНО-СТАТИСТИЧЕСКИЙ АНАЛИЗ

Аннотация. Предложена новая малопараметрическая модель дискретного временного ряда на основе экспоненциального семейства дискретных распределений вероятностей с многомерным параметром. Для параметров предложенной модели строится семейство состоятельных асимптотически нормальных статистических оценок явного вида, в котором найдена асимптотически эффективная оценка, достигающая границы Крамера – Рао при растущей длительности наблюдения временного ряда. Полученные результаты могут быть использованы для робастного статистического анализа дискретных временных рядов, статистического анализа дискретных пространственно-временных данных и случайных полей.

Ключевые слова: дискретный временной ряд, малопараметрическая модель, цепь Маркова высокого порядка, экспоненциальное семейство, эффективная оценка

Для цитирования. Волошко, В. А. Дискретные временные ряды на основе экспоненциального семейства с многомерным параметром и их вероятностно-статистический анализ / В. А. Волошко, Ю. С. Харин // Вес. Нац. акад. навук Беларусі. Сер. фіз.-мат. навук. – 2022. – Т. 58, № 3. – С. 280–291. https://doi.org/10.29235/1561-2430-2022-58-3-280-291

Valeriy A. Voloshko, Yuriy S. Kharin

Research Institute for Applied Problems of Mathematics and Informatics, Belarusian State University, Minsk, Republic of Belarus

DISCRETE-VALUED TIME SERIES BASED ON THE EXPONENTIAL FAMILY WITH THE MULTIDIMENSIONAL PARAMETER AND THEIR PROBABILISTIC AND STATISTICAL ANALYSIS

Abstract. We propose herein a new parsimonious Markov model for a discrete-valued time series with conditional probability distributions of observations lying in the exponential family with the multidimensional parameter. A family of explicit consistent asymptotically normal statistical estimators is constructed for the parameters of the proposed model for increasing length of observed time series, and asymptotically effective estimator is found within this constructed family. The obtained results can be used for robust statistical analysis of discrete-valued time series,and for statistical analysis of discrete-valued spatio-temporal data and random fields.

Keywords: discrete-valued time series, parsimonious model, high-order Markov chain, exponential family, effective estimator

For citation. Voloshko V. A., Kharin Yu. S. Discrete-valued time series based on the exponential family with the multidimensional parameter and their probabilistic and statistical analysis. *Vestsi Natsyianal'nai akademii navuk Belarusi. Seryia fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2022, vol. 58, no. 3, pp. 280–291 (in Russian). https://doi.org/10.29235/1561-2430-2022-58-3-280-291

Введение. Вероятностное моделирование и статистический анализ дискретных временных рядов (ДВР) являются актуальными задачами [1, 2] вследствие, с одной стороны, активной цифровизации всех сфер деятельности и наступления эпохи гиперинформации и, с другой стороны, пока еще недостаточного уровня развития математической теории ДВР в сравнении с непрерывными временными рядами [3, 4].

Один из наиболее развитых и одновременно активно развиваемых подходов к вероятностному моделированию ДВР основан на так называемой обобщенной линейной модели регрессии

© Волошко В. А., Харин Ю. С., 2022

GLM (Generalized Linear Model) [5]. GLM определяет условное математическое ожидание отклика Y при фиксированных факторах X в виде

$$\mu = \mathsf{E}\{Y \mid X\} = F(AX), \tag{1}$$

где A – матрица коэффициентов необходимого формата, $F(\cdot)$ – некоторая функция связи. В случае авторегрессионного ДВР откликом служит наблюдение $Y = x_t$ в момент $t \in \mathbb{Z}$, факторами служат $s \in \mathbb{N}$ предшествующих наблюдений $X = (x_{t-s}, \ldots, x_{t-1})$ (предыстория). В классической модели GLM [5, 6] условное распределение вероятностей $\mathsf{L}\{Y \mid X\}$ лежит в некотором экспоненциальном семействе $\Re$, которое одномерно по параметру (например, биномиальное семейство с фиксированным числом испытаний, или семейство Пуассона); при этом так называемым двойственным параметром [7] служит математическое ожидание $\mu \in \mathbb{R}$ в (1). В [8] предложена модель векторной GLM – обобщение модели (1), где семейство $\Re$ может быть многомерным, а его двойственный параметр может задавать не только математическое ожидание, но и другие численные характеристики отклика (см. пример 4).

В работах [9–12] построены условно нелинейные авторегрессионные модели ДВР, или CNAR-модели (Conditionally Nonlinear AutoRegression), в которых условное распределение вероятностей имеет одномерный параметр. От классических авторегрессионных моделей на основе GLM модели CNAR отличаются тем, что факторами могут быть произвольные нелинейные функции от предыстории, а функция связи $F(\cdot): \mathbb{R} \to (0,1)$ – произвольной абсолютно непрерывной функцией распределения. Отметим, что используемые в GLM-моделях ДВР оценки максимального правдоподобия чувствительны к «тяжелым хвостам» функции распределения $F(\cdot)$, которые ведут к нарушению свойства вогнутости функции лог-правдоподобия [10] и, как следствие, возможности появления у нее множественных локальных максимумов. Такого недостатка лишены построенные для CNAR-моделей новые оценки параметров, называемые FBE (Frequencies-Based Estimators), основанные на частотах, имеющие явный вид и допускающие рекурсивное вычисление при расширении модели.

В настоящей работе предложено обобщение CNAR-модели на случай, когда условное распределение вероятностей является элементом d-мерного экспоненциального семейства.

Модель DCNAR(s,d) и ее свойства. ***Экспоненциальное семейство дискретных распределений вероятностей и его свойства.*** Прежде чем перейти к описанию модели, приведем некоторые базовые свойства экспоненциальных семейств распределений вероятностей [7, 13]. Введем следующие обозначения: $(\Omega, \mathcal{F}, \mathsf{P})$ – вероятностное пространство; $\mathsf{P}\{\cdot\}$, $\mathsf{E}\{\cdot\}$, $\mathsf{D}\{\cdot\}$, $\mathsf{L}\{\cdot\}$, $\mathsf{COV}\{\cdot,\cdot\}$, $\mathsf{I}(\cdot)$ – соответственно функционалы вероятности события, математического ожидания, дисперсии и закона распределения случайной величины, матрицы ковариации пары случайных величин и информационной матрицы Фишера относительно параметра; $z^{\dagger}$ – транспонирование; 0_n, $0_{n\times k}$ – соответственно нулевой n-вектор-столбец и нулевая $(n \times k)$-матрица; $\boldsymbol{X}$ – конечное множество; $\wp = \left\{p(x): \boldsymbol{X} \to [0,1], \sum_{x\in\boldsymbol{X}} p(x) = 1\right\}$ – множество вероятностных распределений на $\boldsymbol{X}$; $\wp^{-} = \wp \setminus \partial\wp = \left\{p(x): \boldsymbol{X} \to \mathbb{R}_{+}, \sum_{x\in\boldsymbol{X}} p(x) = 1\right\}$ – множество невырожденных вероятностных распределений на $\boldsymbol{X}$ (здесь $\partial\wp$ – граница, $\mathbb{R}_{+} = (0, +\infty) \subset \mathbb{R}$ – множество строго положительных действительных чисел); $\Re \subset \wp^{-}$ – d-мерное ($d \le |\boldsymbol{X}| - 1$) экспоненциальное параметрическое семейство дискретных распределений вероятностей на множестве $\boldsymbol{X}$ с d-мерным каноническим параметром $\boldsymbol{\eta} = (\eta_i)_{i=1}^{d} \in \mathbb{R}^{d}$ и функцией вероятностей:

$$\mathsf{P}\{\xi = x\} = \Re(x;\boldsymbol{\eta}) = \frac{1}{Z(\boldsymbol{\eta})}\exp\left(\hbar_0(x) + \left\langle \boldsymbol{\hbar}(x), \boldsymbol{\eta}\right\rangle\right), \quad x \in \boldsymbol{X}, \tag{2}$$

$$\boldsymbol{\hbar}(x) = \left(\hbar_i(x)\right)_{i=1}^{d} \in \mathbb{R}^{d}, \quad \left\langle \boldsymbol{\hbar}(x), \boldsymbol{\eta}\right\rangle = \sum_{i=1}^{d} \hbar_i(x)\eta_i,$$

$$Z(\boldsymbol{\eta}) = \sum_{x\in\boldsymbol{X}} \exp\left(\hbar_0(x) + \left\langle \boldsymbol{\hbar}(x), \boldsymbol{\eta}\right\rangle\right).$$

Здесь $\mathsf{L}\{\xi\} = \Re(\boldsymbol{\eta}) \in \Re$ – закон распределения вероятностей (2); функции $\hbar_i(\cdot): \boldsymbol{X} \to \mathbb{R}$, $i \in \{0, \ldots, d\}$, определяют экспоненциальное семейство $\Re$ и его так называемую [7] каноническую параметризацию; функция $\hbar_0(\cdot)$ может быть произвольной, а функции $\{\hbar_i(\cdot)\}_{i=1}^{d}$ удовлетворяют следующему условию:

$$\langle \hbar(x), c\rangle \neq \text{const}, \quad \forall c \in \mathbb{R}^d \setminus \{0_d\}, \tag{3}$$

другими словами, любая их ненулевая линейная комбинация не может быть константой на $\boldsymbol{X}$. Функция $Z(\boldsymbol{\eta}): \mathbb{R}^d \to \mathbb{R}_+$ является нормировочным коэффициентом в (2). Примем обозначения:

$$\phi(\boldsymbol{\eta}) = \ln Z(\boldsymbol{\eta}), \quad \dot{\phi}(\boldsymbol{\eta}) = \frac{d}{d\boldsymbol{\eta}}\phi(\boldsymbol{\eta}) \in \mathbb{R}^d, \quad \ddot{\phi}(\boldsymbol{\eta}) = \frac{d^2}{d\boldsymbol{\eta}^2}\phi(\boldsymbol{\eta}) \in \mathbb{R}^{d\times d}.$$

Л е м м а 1. *Все распределения вероятностей в семействе $\mathfrak{R}$ невырождены: $\mathfrak{R} \subset \wp^-$.*

Д о к а з а т е л ь с т в о следует из строгой положительности экспонент и нормировочного коэффициента $Z(\boldsymbol{\eta})$ в (2).

Двойственным параметром экспоненциального семейства $\mathfrak{R}$ называется d-вектор [7]:

$$\boldsymbol{\theta} = (\theta_i)_{i=1}^{d} = \dot{\phi}(\boldsymbol{\eta}), \quad \theta_i = \frac{\partial}{\partial \eta_i}\phi(\boldsymbol{\eta}), \quad i \in \{1,\ldots,d\}. \tag{4}$$

Л е м м а 2. *Преобразование* (4) *канонического параметра в двойственный биективно и бесконечно дифференцируемо, а двойственный параметр допускает представление*

$$\boldsymbol{\theta} = \mathsf{E}\{\mathbf{h}(\xi)\}, \quad \theta_i = \mathsf{E}\{\hbar_i(\xi)\}, \quad i \in \{1,\ldots,d\}, \quad \mathsf{L}\{\xi\} = \mathfrak{R}(\boldsymbol{\eta}). \tag{5}$$

Д о к а з а т е л ь с т в о. В силу (4)

$$\theta_i = \frac{\partial}{\partial \eta_i}\ln Z(\boldsymbol{\eta}) = \frac{1}{Z(\boldsymbol{\eta})}\frac{\partial}{\partial \eta_i}Z(\boldsymbol{\eta}),$$

откуда получаем (5). По своему построению отображение $\dot{\phi}(\boldsymbol{\eta})$ бесконечно дифференцируемо, а его биективность следует [7] из невырожденности матрицы производных $\frac{d}{d\boldsymbol{\eta}}\dot{\phi}(\boldsymbol{\eta}) = \ddot{\phi}(\boldsymbol{\eta})$. Действительно, возьмем произвольный ненулевой d-вектор $c \in \mathbb{R}^d \setminus \{0_d\}$ и случайный вектор $\kappa = \hbar(\xi) \in \mathbb{R}^d$ из (5), имеющий математическое ожидание $\mathsf{E}\{\kappa\} = \boldsymbol{\theta}$. Имеем

$$\ddot{\phi}(\boldsymbol{\eta}) = \frac{d}{d\boldsymbol{\eta}}\frac{dZ(\boldsymbol{\eta})/d\boldsymbol{\eta}}{Z(\boldsymbol{\eta})} = \frac{d^2 Z(\boldsymbol{\eta})/d\boldsymbol{\eta}^2}{Z(\boldsymbol{\eta})} - \boldsymbol{\theta}\boldsymbol{\theta}^\dagger = \mathsf{E}\{\kappa\kappa^\dagger\} - \boldsymbol{\theta}\boldsymbol{\theta}^\dagger = \mathsf{COV}\{\kappa,\kappa\},$$

откуда

$$c^\dagger \ddot{\phi}(\boldsymbol{\eta}) c = c^\dagger \mathsf{COV}\{\kappa,\kappa\} c = \mathsf{D}\{\langle c,\kappa\rangle\} > 0.$$

Строгая положительность дисперсии случайной величины $\langle c,\kappa\rangle$ следует из условия (3), по которому эта случайная величина не равна константе.

Таким образом, $c^\dagger \ddot{\phi}(\boldsymbol{\eta}) c > 0$ для любого ненулевого d-вектора c, поэтому симметричная матрица $\ddot{\phi}(\boldsymbol{\eta})$ положительно определена, а значит, невырождена. Лемма 2 доказана.

Из гладкости и биективности отображения $\dot{\phi}(\boldsymbol{\eta})$ следует его непрерывная обратимость, поэтому близким каноническим параметрам $\boldsymbol{\eta}, \tilde{\boldsymbol{\eta}}$ отвечают близкие двойственные параметры $\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}$ и наоборот. Задача восстановления канонического параметра $\boldsymbol{\eta}$ по двойственному параметру $\boldsymbol{\theta}$ эквивалентна обращению функции $\dot{\phi}(\cdot)$, и в общем случае обратная функция $\dot{\phi}^{-1}(\cdot)$ не имеет явного вида. Однако эта задача также эквивалентна задаче максимизации:

$$L(z) \to \max_{z\in\mathbb{R}^d}, \quad L(z) = \langle z, \boldsymbol{\theta}\rangle - \phi(z), \quad z \in \mathbb{R}^d.$$

Функция $L(z)$ строго вогнута, поскольку $\frac{d^2}{dz^2}L(z) = -\ddot{\phi}(z) < 0$, и из уравнения $0_d = \frac{d}{dz}L(z) = \boldsymbol{\theta} - \dot{\phi}(z)$ выражается ее единственный локальный максимум $z = \boldsymbol{\eta}$, который может быть приближенно вычислен численным методом градиентного подъема.

Для удобства в случае задания представителя семейства $\Re$ через двойственный параметр $\boldsymbol{\theta}$ будем использовать обозначения, аналогичные случаю задания через канонический параметр $\boldsymbol{\eta}$, но с заменой круглых скобок на квадратные, т. е. положим $\Re[\boldsymbol{\theta}] = \Re(\boldsymbol{\eta})$ и $\Re[x;\boldsymbol{\theta}] = \Re(x;\boldsymbol{\eta})$. Введем следующие множества:

$$\boldsymbol{\Theta} = \mathrm{conv}\{\mathbf{h}(x) : x \in \boldsymbol{X}\}, \quad \boldsymbol{\Theta}^- = \boldsymbol{\Theta} \setminus \partial\boldsymbol{\Theta},$$

где $\mathrm{conv}\{\cdot\}$ – выпуклая оболочка множества. Выпуклый многогранник $\boldsymbol{\Theta}$ – область значений векторов математических ожиданий $\mathsf{E}\{\hbar(\xi)\}$ для всех случайных величин $\xi \in \boldsymbol{X}$, $\mathsf{L}\{\xi\} \in \wp$, внутренность $\boldsymbol{\Theta}^-$ многогранника $\boldsymbol{\Theta}$ – область значений двойственного параметра $\boldsymbol{\theta}$ [7], совпадающая с областью значений векторов $\mathsf{E}\{\hbar(\xi)\}$ для всех случайных величин $\xi \in \boldsymbol{X}$, имеющих невырожденное распределение вероятностей $\mathsf{L}\{\xi\} \in \wp^-$. Пусть наблюдается последовательность $n \in \mathbb{N}$ независимых одинаково распределенных случайных величин $\xi_1,\ldots,\xi_n \in \boldsymbol{X}$, $\mathsf{L}\{\xi\} \equiv \Re[\boldsymbol{\theta}]$. Тогда оценка максимального правдоподобия (ОМП) двойственного параметра $\boldsymbol{\theta}$ имеет вид [7]

$$\hat{\boldsymbol{\theta}} = \frac{1}{n}\sum_{i=1}^{n} \hbar(\xi_i). \tag{6}$$

При этом ОМП канонического параметра $\boldsymbol{\eta}$ существует при условии $\hat{\boldsymbol{\theta}} \in \boldsymbol{\Theta}^-$ и имеет вид $\hat{\boldsymbol{\eta}} = \dot{\phi}^{-1}(\hat{\boldsymbol{\theta}})$ (в случае $\hat{\boldsymbol{\theta}} \in \partial\boldsymbol{\Theta}$ функция правдоподобия не имеет глобального максимума на $\mathbb{R}^d$).

ДВР на основе экспоненциального семейства с многомерным параметром и его частные случаи. Пусть на вероятностном пространстве $(\Omega, \mathsf{F}, \mathsf{P})$ определен ДВР $x_t \in \boldsymbol{X}$, $t \in \mathbb{Z} = \{\ldots,-1,0,+1,\ldots\}$, задаваемый соотношениями

$$\mathsf{L}\{x_t \mid x_\tau : \tau < t\} = \Re\left[\boldsymbol{\theta}^{X_{t-s}^{t-1}}\right] \equiv \Re\left[\boldsymbol{\eta}^{X_{t-s}^{t-1}}\right], \tag{7}$$

$$\boldsymbol{\theta}^{X_{t-s}^{t-1}} = \left(\theta_i^{X_{t-s}^{t-1}}\right)_{i=1}^{d} = F\left(\sum_{j=1}^{m} \zeta_j \psi^j\left(X_{t-s}^{t-1}\right)\right) \in \mathbb{R}^d, \tag{8}$$

$$\boldsymbol{\eta}^{X_{t-s}^{t-1}} = \left(\eta_i^{X_{t-s}^{t-1}}\right)_{i=1}^{d} = \dot{\phi}^{-1}\left(\boldsymbol{\theta}^{X_{t-s}^{t-1}}\right) = \tilde{F}\left(\sum_{j=1}^{m} \zeta_j \psi^j\left(X_{t-s}^{t-1}\right)\right) \in \mathbb{R}^d, \tag{9}$$

$$\tilde{F} = \dot{\phi}^{-1} \circ F, \tag{10}$$

где «$\circ$» – знак композиции функций, $\zeta_j \in \mathbb{R}$, $j \in \{1,\ldots,m\}$, – параметры модели (вектор параметров модели будем обозначать $\boldsymbol{\zeta} = (\zeta_j)_{j=1}^m \in \mathbb{R}^m$), $X_{t-s}^{t-1} = (x_{t-s},\ldots,x_{t-1}) \in \boldsymbol{X}^s$ – предыстория глубины $s \in \mathbb{N}$ (s-предыстория) наблюдения x_t, $\psi^j(\cdot) = \left(\psi_i^j(\cdot)\right)_{i=1}^{d} : \boldsymbol{X}^s \to \mathbb{R}^d$, $j \in \{1,\ldots,m\}$, – линейно независимые на $\boldsymbol{X}^s$ базисные функции, $F(\cdot) : \mathbb{R}^d \to \boldsymbol{\Theta}^-$ – дважды непрерывно дифференцируемое взаимно-однозначное отображение (будем называть это условие **условием регулярности A.1**), $\dot{\phi}^{-1}$ обозначает функцию, обратную к $\dot{\phi}$. Соотношение (7) означает, что условное распределение наблюдения x_t при фиксированной предыстории $\{x_\tau : \tau < t\}$ зависит только от s-предыстории X_{t-s}^{t-1} и лежит в экспоненциальном семействе $\Re$. Соотношение (8) определяет двойственный параметр $\boldsymbol{\theta}^{X_{t-s}^{t-1}} \in \boldsymbol{\Theta}^-$ условного распределения $\mathsf{L}\{x_t \mid x_\tau : \tau < t\}$ как функцию от s-предыстории X_{t-s}^{t-1} с параметром $\boldsymbol{\zeta}$. Соотношение (9) эквивалентно соотношению (8) и определяет канонический параметр $\boldsymbol{\eta}^{X_{t-s}^{t-1}} \in \mathbb{R}^d$ условного распределения $\mathsf{L}\{x_t \mid x_\tau : \tau < t\}$.

Основной материал статьи будем излагать в терминах двойственных, а не канонических параметров условных распределений, что связано с более простым видом ОМП (6) в сравнении с ОМП $\hat{\boldsymbol{\eta}} = \dot{\phi}^{-1}(\hat{\boldsymbol{\theta}})$. Поэтому из двух эквивалентных соотношений (8), (9) при изложении основных результатов будем использовать соотношение (8). Понятие канонической функции связи и смежные вопросы будут рассмотрены в терминах канонических параметров условных распределений и определяемого (10) дважды непрерывно дифференцируемого взаимно-однозначного отображения $\tilde{F} : \mathbb{R}^d \to \mathbb{R}^d$ (эти свойства эквивалентны условию регулярности **A.1**). Модель (7),

(8) будем называть дискретной условно нелинейной авторегрессией порядка s с d-мерным параметром $(1 \le d \le |\boldsymbol{X}| - 1)$ DCNAR(s,d) (Discrete Conditionally Nonlinear AutoRegression). Если нужно указать используемое экспоненциальное семейство $\Re$, будем к названию модели добавлять название семейства $\Re$, а к аббревиатуре CNAR – соответствующий префикс, например, мультиномиальная условно нелинейная авторегрессия MultCNAR.

Модель DCNAR(s,d), определяемую (7), (8), будем называть канонической, если $F(\cdot) = \dot{\phi}(\cdot)$, или эквивалентно, если функция $\tilde{F}$, определяемая (10), является тождественным преобразованием: $\tilde{F}(z) = z,\ z \in \mathbb{R}^d$.

Для канонической модели DCNAR(s,d), согласно (9), канонический параметр условного распределения $\mathsf{L}\{x_t \mid x_\tau : \tau < t\}$ задается напрямую (без дополнительного преобразования) линейной комбинацией базисных функций $\{\psi^j(\cdot)\}_{j=1}^m$ от предыстории X_{t-s}^{t-1}.

Приведем примеры – частные случаи моделей CNAR(s,d). Всюду в примерах $N \in \mathbb{N}$,

$$\Lambda(z) = \frac{1}{1+e^{-z}}, \quad z \in \mathbb{R}, \tag{11}$$

– логистическая функция распределения (ф. р.), $\Lambda^{-1}(z) = \ln\left(\dfrac{z}{1-z}\right)$, $z \in (0,1)$, – соответствующая квантильная функция.

П р и м е р 1 [11, 12, 14]. Пусть $d = 1$, $\boldsymbol{X} = \{0,\ldots,N\}$, $\hbar_0(x) = \ln\binom{N}{x}$, $\hbar_1(x) = x$, $\boldsymbol{\Theta} = [0,N]$, $\boldsymbol{\Theta}^- = (0,N)$, $\phi(\boldsymbol{\eta}) = N\ln(1+e^{\boldsymbol{\eta}})$, $\boldsymbol{\theta} = \dot{\phi}(\boldsymbol{\eta}) = N\Lambda(\boldsymbol{\eta})$, $\boldsymbol{\eta} = \dot{\phi}^{-1}(\boldsymbol{\theta}) = \Lambda^{-1}(\boldsymbol{\theta}/N)$. В этом частном случае модель CNAR(s,d) задает биномиальную условно нелинейную авторегрессию BiCNAR(s) порядка s [11, 14], а при $N = 1$, ее частный случай, – двоичную условно нелинейную авторегрессию BCNAR(s) [12]. Стандартный параметр биномиального распределения – «вероятность успеха» – во введенных терминах равен $\boldsymbol{\theta}/N \in (0,1)$. Если функцию $\hbar_0(\cdot)$ сделать произвольной – получим семибиномиальную условно нелинейную авторегрессию, исследованную в [10].

П р и м е р 2 [8]. Пусть $d \in \mathbb{N}$, $\boldsymbol{X} = \{0,\ldots,N\}^d = \left\{x = (x_i)_{i=1}^d : x_i \in \{0,\ldots,N\}\right\}$, $\hbar_0(x) = \sum_{i=1}^d \ln\binom{N}{x_i}$, $\hbar_i(x) = x_i,\ \ i \in \{1,\ldots,d\}$, $\boldsymbol{\Theta} = [0,N]^d$, $\boldsymbol{\Theta}^- = (0,N)^d$, $\phi(\boldsymbol{\eta}) = N\sum_{i=1}^d \ln(1+e^{\eta_i})$, $\theta_i = \dfrac{\partial}{\partial \eta_i}\phi(\boldsymbol{\eta}) = N\Lambda(\eta_i)$, $\eta_i = \Lambda^{-1}(\theta_i/N),\ \ i \in \{1,\ldots,d\}$. В этом частном случае модель CNAR(s,d) задает d-мерный векторный аналог модели BiCNAR(s): $x_t = (x_{t,i})_{i=1}^d$, где компоненты d-вектора x_t при фиксированной предыстории $\{x_\tau : \tau < t\}$ условно независимы и имеют (в общем случае различные) биномиальные условные распределения вероятностей $\mathsf{L}\{x_{t,i} \mid x_\tau : \tau < t\} = \mathrm{Bi}(N, \theta_i/N)$ (отметим, что условное распределение компоненты $x_{t,i}$, вообще говоря, зависит от всей матрицы компонент s-предыстории $(x_{t-j,k})_{j\in\{1,\ldots,s\},k\in\{1,\ldots,d\}}$, а не только от вектора значений той же компоненты в прошлом $(x_{t-j,i})_{j\in\{1,\ldots,s\}}$, поэтому векторный временной ряд $\{x_t\}$ в общем случае не распадается на d независимых покомпонентных временных рядов $\{x_{t,i}\},\ i \in \{1,\ldots,d\}$).

П р и м е р 3 [15, 16]. Пусть $d \in \mathbb{N}$, $\boldsymbol{X} = \left\{x = (x_i)_{i=0}^d \in \mathbb{N}_0^{d+1} : \sum_{i=0}^d x_i = N\right\}$, $\hbar_0(x) = \ln\binom{N}{x_0,\ldots,x_d}$, $\hbar_i(x) = x_i,\ \ i \in \{1,\ldots,d\}$, $\boldsymbol{\Theta}$ – симплекс с вершинами 0_d и $N \cdot \boldsymbol{e}^i$, $i \in \{1,\ldots,d\}$, где $\boldsymbol{e}^i = (\delta_{i,j})_{j=1}^d$ – канонический базис в $\mathbb{R}^d$, $\delta_{i,j}$ – символ Кронекера, $\phi(\boldsymbol{\eta}) = N\ln\left(1+\sum_{j=1}^d e^{\eta_j}\right)$, $\theta_i = \dfrac{\partial}{\partial \eta_i}\phi(\boldsymbol{\eta}) = Ne^{\eta_i}\left(1+\sum_{j=1}^d e^{\eta_j}\right)^{-1}$, $\eta_i = \ln\left(\dfrac{\theta_i}{N-\sum_{j=1}^d \theta_j}\right)$. В этом случае модель CNAR(s,d) задает мультиномиальный аналог MultCNAR(s,d) биномиальной модели BiCNAR(s): $x_t = (x_{t,i})_{i=0}^d$ (при $d = 1$ последовательность $y_t = x_{t,1}$ описывается моделью BiCNAR(s)), где вектор x_t при фиксированной предыстории $\{x_\tau : \tau < t\}$ условно мультиномиально распределен: $\mathsf{L}\{x_t \mid x_\tau : \tau < t\} = \mathrm{Mult}(N, p_0,\ldots,p_d)$, $p_0 = 1 - \sum_{i=1}^d p_i$, $p_i = \theta_i^{X_{t-s}^{t-1}}/N,\ \ i \in \{1,\ldots,d\}$.

П р и м е р 4. Пусть $d = 2 \le N$, $\boldsymbol{X} = \{0,\ldots,N\}$, $\hbar_0(x) = \ln\binom{N}{x}$, $\hbar_1(x) = x$, $\hbar_2(x) = x^2$, $\phi(\boldsymbol{\eta}) = \ln\sum_{x\in\boldsymbol{X}}\binom{N}{x}e^{\eta_1 x + \eta_2 x^2}$, $\boldsymbol{\Theta} \subset \mathbb{R}^2$ – плоский $(N+1)$-угольник с вершинами (x,x^2), $x \in \boldsymbol{X}$. Определим множество $\tilde{\boldsymbol{\Theta}}^-$ допустимых значений пары $(\mu,\sigma^2) = (\mathsf{E}\{\xi\},\mathsf{D}\{\xi\})$ (математическое ожидание и дисперсия) для случайных величин $\xi \in \boldsymbol{X}$, имеющих невырожденное распределение вероятностей $\mathsf{L}\{\xi\} \in \wp^-$:

$$\tilde{\boldsymbol{\Theta}}^- ::= \rho(\boldsymbol{\Theta}^-), \quad \rho(\boldsymbol{\theta}) ::= (\theta_1, \theta_2 - \theta_1^2) = (\mu, \sigma^2), \quad \boldsymbol{\theta} = (\theta_1, \theta_2) \in \boldsymbol{\Theta}^-, \tag{12}$$

где запрещены только граничные значения, например такие, как $(\mu, \sigma^2) = (0,0)$. В этом случае модель CNAR(s,2) задает аналог биномиальной модели BiCNAR(s) через расширение семейства $\Re$ условных переходных распределений вероятностей $\mathsf{L}\{x_t \mid x_\tau : \tau < t\}$. Экспоненциальное семейство $\Re$ содержит в себе биномиальное семейство ($\Re(\boldsymbol{\eta}) \in \mathrm{Bi}$ при $\eta_2 = 0$), и для любой допустимой пары $(\mu, \sigma^2) \in \tilde{\boldsymbol{\Theta}}^-$ существует единственное распределение вероятностей в $\Re$, имеющее математическое ожидание μ и дисперсию σ^2. Двойственный параметр такого распределения, согласно (12), имеет вид

$$\boldsymbol{\theta} = \rho^{-1}(\mu, \sigma^2) = (\mu, \mu^2 + \sigma^2). \tag{13}$$

Описанная модель CNAR(s,2) с помощью (13) позволяет задавать не только математическое ожидание, но и дисперсию условного распределения следующего наблюдения как функцию от предыстории. Функция связи канонического и двойственного параметров

$$\boldsymbol{\theta} = \dot{\phi}(\boldsymbol{\eta}) = \left(\sum_{x \in \boldsymbol{X}} \binom{N}{x} x e^{\eta_1 x + \eta_2 x^2}, \sum_{x \in \boldsymbol{X}} \binom{N}{x} x^2 e^{\eta_1 x + \eta_2 x^2} \right) \Big/ \sum_{x \in \boldsymbol{X}} \binom{N}{x} e^{\eta_1 x + \eta_2 x^2}$$

обращается численным методом градиентного подъема.

Теорема 1. *При условии регулярности* **A.1** *ДВР* $\{x_t\}$, *определяемый моделью* (7), (8), *представляет собой эргодическую цепь Маркова порядка s с положительными переходными вероятностями*

$$p_{q \to x} ::= \mathsf{P}\{x_t = x \mid X_{t-s}^{t-1} = q\} = \Re[x; \boldsymbol{\theta}^q] > 0, \quad x \in \boldsymbol{X}, \quad q \in \boldsymbol{X}^s,$$

и имеет единственное невырожденное стационарное распределение вероятностей $\pi = (\pi_q)_{q \in \boldsymbol{X}^s}$, $\pi_q = \mathsf{P}\{X_{t-s}^{t-1} = q\} > 0$, $q \in \boldsymbol{X}^s$.

Доказательство идентично доказательству теоремы 1 в [9] и основано на строгой положительности всех переходных вероятностей $p_{q \to x}$, которая следует из леммы 1.

В случае d = 1 и, без потери общности, при $\boldsymbol{\Theta}^- = (0,1)$ в качестве функций $F(\cdot): \mathbb{R} \to (0,1)$ используются абсолютно непрерывные ф. р. [5, 6], среди которых выделяют ф. р. с «легкими» и «тяжелыми хвостами». Наиболее распространенные представители этих двух типов – соответственно ф. р. Гаусса (легкие хвосты) и ф. р. Коши (тяжелые хвосты):

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt, \quad C(z) = \frac{1}{2} + \frac{\mathrm{arctg}\,(z)}{\pi}, \quad z \in \mathbb{R}. \tag{14}$$

Канонические ф. р., такие как логистическая ф. р. (11) (см. примеры 1, 2), имеют экспоненциально убывающие хвосты, которые относятся к промежуточному типу. Пусть в (10) $\dot{\phi}(\cdot) = \Lambda(\cdot)$ и $F(\cdot)$ – одна из двух ф. р. (14). Рассмотрим асимптотику соответствующих двух функций $\tilde{F}(z): \mathbb{R} \to \mathbb{R}$ при $z \to +\infty$ (асимптотика при $z \to -\infty$ получается аналогично ввиду нечетности обеих функций $\tilde{F}(\cdot)$). Для ф. р. Гаусса $F(\cdot) = \Phi(\cdot)$ имеем

$$\tilde{F}(z) = \Lambda^{-1}\left(\Phi(z)\right) = \ln\left(\frac{\Phi(z)}{1 - \Phi(z)} \right) \underset{z \to +\infty}{\sim} -\ln\left(1 - \Phi(z)\right) \sim \frac{z^2}{2}.$$

Для ф. р. Коши $F(\cdot) = C(\cdot)$ аналогично

$$\tilde{F}(z) \sim -\ln\left(1 - C(z)\right) \sim -\ln\left(\frac{1}{2} - \frac{\pi/2 - \mathrm{arctg}\,(1/z)}{\pi} \right) \sim \ln z.$$

Другими словами, при $|z| \to +\infty$ легкие хвосты у $F(\cdot)$ соответствуют асимптотике $|\tilde{F}(z)|/|z| \to +\infty$ у $\tilde{F}(\cdot)$, тяжелые хвосты у $F(\cdot)$ соответствуют асимптотике $|\tilde{F}(z)|/|z| \to 0$ у $\tilde{F}(\cdot)$. По аналогии мы можем обобщить понятия легких и тяжелых хвостов на случай произвольных $d \in \mathbb{N}$ и функций $F(\cdot): \mathbb{R}^d \to \boldsymbol{\Theta}^-$. А именно, аналогом свойства легких (тяжелых) хвостов у функции $F(\cdot)$ можно считать асимптотику $\|\tilde{F}(z)\| / \|z\| \to +\infty$ $(\to 0)$ при $\|z\| \to +\infty$ у отвечающей ей функции $\tilde{F}(\cdot)$, где $\|\cdot\|$ – некоторая норма в $\mathbb{R}^d$. Мерой легкости (тяжести) хвостов можно считать скорость роста (убывания) отношения $\|\tilde{F}(z)\| / \|z\|$ при $\|z\| \to +\infty$.

Статистические оценки параметров модели DCNAR(*s,d*). Пусть наблюдается ДВР $x_1, \ldots, x_T \in \boldsymbol{X}$ длины $T > s$, определяемый моделью (7), (8). Для оценивания параметров этой модели DCNAR(*s,d*) применим FBE-метод [10, 12, 14].

Введем обозначения: $\mathbf{1}\{\cdot\}$ – индикаторная функция;

$$\hat{\pi}_q = \frac{1}{T-s} \sum_{t=s+1}^{T} \mathbf{1}\left\{ X_{t-s}^{t-1} = q \right\}, \quad \boldsymbol{M}_T = \left\{ q \in \boldsymbol{X}^s : \hat{\pi}_q > 0 \right\} \subset \boldsymbol{X}^s \tag{15}$$

– соответственно статистическая оценка *s*-мерного стационарного распределения и ее носитель;

$$\hat{\boldsymbol{\theta}}^q = \frac{\sum_{t=s+1}^{T} \hbar(x_t) \cdot \mathbf{1}\left\{ X_{t-s}^{t-1} = q \right\}}{\sum_{t=s+1}^{T} \mathbf{1}\left\{ X_{t-s}^{t-1} = q \right\}}, \quad \boldsymbol{M}_T^+ = \left\{ q \in \boldsymbol{M}_T : \hat{\boldsymbol{\theta}}^q > 0 \right\} \subset \boldsymbol{M}_T \tag{16}$$

– соответственно статистическая оценка двойственного параметра $\boldsymbol{\theta}^q$ условного распределения $\mathsf{L}\{x_t \mid X_{t-s}^{t-1} = q\}$, согласно лемме 2, и подмножество предысторий $q \in \boldsymbol{X}^s$, для которых $\hat{\boldsymbol{\theta}}^q$ не лежит на границе многогранника $\boldsymbol{\Theta}$. Отметим, что оценка (16) является ОМП для параметров модели (7) (это доказывается аналогично леммам 3 и 4 в [10] и следует из общего вида (6) ОМП двойственного параметра экспоненциального семейства).

Л е м м а 3. *При условии регулярности* **A.1** *множество* (15) *удовлетворяет асимптотике*

$$\boldsymbol{M}_T \xrightarrow[T \to \infty]{\text{п.н.}} \boldsymbol{X}^s. \tag{17}$$

Д о к а з а т е л ь с т в о идентично доказательству леммы 1 в [9] и основано на теореме 1, состоятельности оценки (15) стационарного распределения для эргодической цепи Маркова $\{x_t\}$ и монотонном неубывании $\boldsymbol{M}_T$ по T.

Введем множество $\boldsymbol{S} = \boldsymbol{X}^s \times \{1, \ldots, d\}$ мощности $|\boldsymbol{S}| = d \cdot |\boldsymbol{X}|^s$ и упорядочим его элементы $(q,i) \in \boldsymbol{S}$ лексикографически так, что элементы $(q,1), \ldots, (q,d)$ идут непосредственно друг за другом при каждом $q \in \boldsymbol{X}^s$. Обозначим: $G(\cdot) = F^{-1}(\cdot): \boldsymbol{\Theta}^- \to \mathbb{R}^d$ – функция связи; если $z = (z_i)_{i=1}^n \in (\boldsymbol{\Theta}^-)^n$, то $G(z) ::= \left(G(z_i) \right)_{i=1}^n \in \mathbb{R}^{nd}$; $\Theta = (\theta_i^q)_{(q,i) \in \boldsymbol{S}} \in \mathbb{R}^{|\boldsymbol{S}|}$ – вектор-столбец, составленный из $|\boldsymbol{X}|^s$ вектор-столбцов $\boldsymbol{\theta}^q \in \mathbb{R}^d$, $q \in \boldsymbol{X}^s$ (аналогично вектор-столбец $\hat{\Theta}$ составлен из определяемых (16) $|\boldsymbol{M}_T^+|$ вектор-столбцов $\hat{\boldsymbol{\theta}}^q \in \boldsymbol{\Theta}^-$, $q \in \boldsymbol{M}_T^+$); $\Gamma = G(\Theta) \in \mathbb{R}^{|\boldsymbol{S}|}$ – вектор-столбец, составленный из вектор-столбцов

$$\boldsymbol{\gamma}^q = \left(\gamma_i^q \right)_{i=1}^d ::= G(\boldsymbol{\theta}^q) = \sum_{i=1}^{m} \zeta_i \psi^i(q), \quad q \in \boldsymbol{X}^s, \tag{18}$$

где соотношение (18) следует из (8); аналогично вектор-столбец $\hat{\Gamma} = G(\hat{\Theta})$ составлен из $|\boldsymbol{M}_T^+|$ вектор-столбцов

$$\hat{\boldsymbol{\gamma}}^q = \left(\hat{\gamma}_i^q \right)_{i=1}^d ::= G(\hat{\boldsymbol{\theta}}^q) \in \mathbb{R}^d, \quad q \in \boldsymbol{M}_T^+;$$

$H = H^\dagger = (h_{(q,i),(q',i')}) \in \mathbb{R}^{|\boldsymbol{S}| \times |\boldsymbol{S}|}$ – некоторая симметричная неотрицательно определенная весовая матрица, $H_{q,q'} = (h_{(q,i),(q',i')})_{i,i'=1}^d \in \mathbb{R}^{d \times d}$ – $(d \times d)$-блок матрицы H для $q, q' \in \boldsymbol{X}^s$; $\Psi = \left(\psi_i^j(q) \right)_{(q,i) \in \boldsymbol{S}, j \in \{1, \ldots, m\}} \in \mathbb{R}^{|\boldsymbol{S}| \times m}$ – матрица, составленная из компонент базисных функций, так что j-й столбец матрицы Ψ составлен из вектор-столбцов $\psi^j(q) = \left(\psi_i^j(q) \right)_{i=1}^d \in \mathbb{R}^d$, $q \in \boldsymbol{X}^s$; $\langle \Psi \rangle_H = \Psi^\dagger H \Psi \in \mathbb{R}^{m \times m}$ – матрица Грама базисных функций $\{\psi^j\}_{j=1}^m$ относительно скалярного произведения:

$$\langle u,v\rangle_H = \sum_{(q,i),(q',i')\in \boldsymbol{S}} u_{q,i} v_{q',i'} h_{(q,i),(q',i')}, \quad u,v:\boldsymbol{S}\to\mathbb{R};$$

для некоторого подмножества s-предысторий $M\subset \boldsymbol{X}^s$, $\Theta(M)$ – вектор-столбец, составленный из вектор-столбцов $\boldsymbol{\theta}^q\in\mathbb{R}^d$, $q\in M$; $H(M)$ – подматрица матрицы H, составленная из блоков $H_{q,q'}$, $q,q'\in M$; $\Psi(M)$ – матрица из m столбцов, где j-й столбец состоит из вектор-столбцов $\psi^j(q)\in\mathbb{R}^d$, $q\in M$.

На основе соотношения (18) и оценок (16) составим систему из $|\boldsymbol{M}_T^+|$ линейных относительно параметра $\boldsymbol{\zeta}\in\mathbb{R}^m$ векторных уравнений:

$$\boldsymbol{\gamma}^q = \hat{\boldsymbol{\gamma}}^q, \quad q\in \boldsymbol{M}_T^+. \tag{19}$$

Каждое векторное уравнение (19) эквивалентно d скалярным уравнениям $\gamma_i^q=\hat{\gamma}_i^q$, $i\in\{1,\ldots,d\}$, поэтому в системе (19) всего $d\cdot|\boldsymbol{M}_T^+|$ скалярных линейных уравнений относительно вектора $\boldsymbol{\zeta}=(\zeta_i)$ с m неизвестными. Как будет показано далее (см. следствие 1), для множества $\boldsymbol{M}_T^+$, определяемого (16), при $T\to\infty$ имеет место почти наверное сходимость вида (17) к полному множеству s-предысторий $\boldsymbol{X}^s$, поэтому будем предполагать, что

$$m < d\cdot|\boldsymbol{M}_T^+| < d\cdot|\boldsymbol{X}|^s.$$

При таком предположении система (19) переопределена, и для ее решения будем минимизировать квадратическую функцию уклонений левой и правой частей, параметризованную весовой матрицей H:

$$W(\boldsymbol{\zeta}) = \sum_{q,q'\in \boldsymbol{M}_T^+} (\boldsymbol{\gamma}^q-\hat{\boldsymbol{\gamma}}^q)^\dagger H_{q,q'}(\boldsymbol{\gamma}^q-\hat{\boldsymbol{\gamma}}^q)\ge 0. \tag{20}$$

Введем дополнительное **условие регулярности** – **A.2**:

$$\det\left(\left\langle \Psi(\boldsymbol{M}_T^+)\right\rangle_{H(\boldsymbol{M}_T^+)}\right)>0.$$

Функция потерь (20) может быть представлена в матричном виде:

$$W(\boldsymbol{\zeta})=\left(\Gamma(\boldsymbol{M}_T^+)-\hat{\Gamma}\right)^\dagger H(\boldsymbol{M}_T^+)\left(\Gamma(\boldsymbol{M}_T^+)-\hat{\Gamma}\right)=\left(\Psi(\boldsymbol{M}_T^+)\boldsymbol{\zeta}-\hat{\Gamma}\right)^\dagger H(\boldsymbol{M}_T^+)\left(\Psi(\boldsymbol{M}_T^+)\boldsymbol{\zeta}-\hat{\Gamma}\right)$$

и при условии регулярности **A.2** имеет единственный минимум в точке:

$$\hat{\boldsymbol{\zeta}} ::= \arg\min_{\boldsymbol{\zeta}\in\mathbb{R}^m} W(\boldsymbol{\zeta}) = \left\langle \Psi(\boldsymbol{M}_T^+)\right\rangle_{H(\boldsymbol{M}_T^+)}^{-1} (\Psi(\boldsymbol{M}_T^+))^\dagger\cdot H(\boldsymbol{M}_T^+)\cdot\hat{\Gamma}. \tag{21}$$

Формула (21) определяет параметрическое семейство (с параметром H) статистических оценок $\hat{\boldsymbol{\zeta}}$ на основе частот параметра $\boldsymbol{\zeta}$ модели DCNAR(s,d). Заметим, что FBE (21) имеет явный вид. Условие регулярности **A.2** не является жестким: в общем случае при строго положительно определенной матрице H условие **A.2** эквивалентно линейной независимости базисных функций $\{\psi^j\}_{j=1}^m$ на $\boldsymbol{M}_T^+\subset\boldsymbol{X}^s$, что асимптотически выполняется ввиду линейной независимости базисных функций на $\boldsymbol{X}^s$ и следствия 1.

Свойства статистических оценок параметров модели DCNAR(s,d).

Т е о р е м а 2. *При условии регулярности* **A.1** *и* $T\to+\infty$ *статистическая оценка* (16) *состоятельна, асимптотически нормальна и асимптотически эффективна (асимптотически достигает границы Крамера – Рао):*

$$\hat{\Theta}\xrightarrow[T\to\infty]{\mathsf{P}}\Theta, \quad \sqrt{T}(\hat{\Theta}-\Theta)\xrightarrow[T\to\infty]{D}\mathbf{N}_{|\boldsymbol{S}|}\left(0_{|\boldsymbol{S}|},\mathsf{I}^{-1}(\Theta)\right),$$

$$\mathsf{I}(\Theta)=\operatorname{diag}\left(\pi_q\mathsf{I}(\boldsymbol{\theta}^q)\right)_{q\in\boldsymbol{X}^s}, \tag{22}$$

где $\mathrm{I}(\Theta)$ *– информационная матрица Фишера параметра* Θ *модели* (7)*,* $\mathrm{diag}(\cdot)$ *– блочно-диагональная матрица,* $\mathrm{I}(\boldsymbol{\theta}) \in \mathbb{R}^{d\times d}$ *– информационная матрица Фишера параметра* $\boldsymbol{\theta}$ *семейства* $\mathfrak{R}$.

Схема доказательства аналогична теореме 3 в [10] и основана на применении теоремы 4.1 в [17], согласно которой состоятельность, асимптотическая нормальность и эффективность ОМП $\hat{\Theta}$ параметра Θ цепи Маркова $\{x_t\}$ следуют из выполнения некоторых условий регулярности на переходные вероятности, обеспеченных условием **A.1**.

Следствие 1. *При условии регулярности* **A.1** *множество* (16) *удовлетворяет асимптотике*

$$\boldsymbol{M}_T^+ \xrightarrow[T\to\infty]{\text{п.н.}} \boldsymbol{X}^s. \tag{23}$$

Доказательство. Сходимость (23) следует из монотонного неубывания $\boldsymbol{M}_T^+$ по T, условий $\boldsymbol{\theta}^q \in \boldsymbol{\Theta}^-$, $q \in \boldsymbol{X}^s$, и теоремы 2. Что и требовалось доказать.

Из (23) следует, что число уравнений в системе (19) асимптотически стремится к максимальному значению $|\boldsymbol{S}| = d\cdot|\boldsymbol{X}|^s$.

Следствие 2. *В обозначениях теоремы* 2*, при выполнении условия регулярности* **A.1** *подстановочная оценка* $\hat{\Gamma}$ *вектора* Γ *состоятельна, асимптотически нормальна и асимптотически эффективна:*

$$\hat{\Gamma} \xrightarrow[T\to\infty]{\mathrm{P}} \Gamma, \quad \sqrt{T}(\hat{\Gamma}-\Gamma) \xrightarrow[T\to\infty]{D} \mathbf{N}_{|\boldsymbol{S}|}\left(0_{|\boldsymbol{S}|}, \mathrm{I}^{-1}(\Gamma)\right),$$

$$\mathrm{I}(\Gamma) = (\boldsymbol{g}^{-1})^{\dagger}\mathrm{I}(\Theta)\boldsymbol{g}^{-1} =:: \boldsymbol{J}, \quad \boldsymbol{g} = \mathrm{diag}\left(\frac{dG(\theta^q)}{d\theta^q}\right)_{q\in\boldsymbol{X}^s}, \tag{24}$$

где $\mathrm{I}(\Gamma)$ *– информационная матрица Фишера параметра* Γ *модели* (7).

Схема доказательства аналогична теореме 2 и использует то свойство, что $\Gamma = G(\Theta)$ – гладкая замена параметра Θ модели (7), и $\hat{\Gamma} = G(\hat{\Theta})$ – ОМП.

Теорема 3. *При условиях регулярности* **A.1***,* **A.2** *оценка* (21) *состоятельна и асимптотически нормальна:*

$$\hat{\boldsymbol{\zeta}} \xrightarrow[T\to\infty]{\mathrm{P}} \boldsymbol{\zeta}, \quad \sqrt{T}(\hat{\boldsymbol{\zeta}}-\boldsymbol{\zeta}) \xrightarrow[T\to\infty]{D} \mathbf{N}_m(0_m, \Sigma_H),$$

$$\Sigma_H = \langle\Psi\rangle_H^{-1}\langle\Psi\rangle_{H\boldsymbol{J}^{-1}H}\langle\Psi\rangle_H^{-1}, \tag{25}$$

где матрица $\boldsymbol{J}$ *определена* (24).

Доказательство идентично доказательству теоремы 3 в [9] и опирается на следствие 2 и то свойство, что оценка (21) – линейное преобразование оценки $\hat{\Gamma}$.

Следствие 3. *Если весовая матрица совпадает с матрицей* (24)

$$H = \boldsymbol{J}, \tag{26}$$

то оценка (21) *асимптотически эффективна:*

$$\Sigma_{\boldsymbol{J}} = \mathrm{I}^{-1}(\boldsymbol{\zeta}), \quad \mathrm{I}(\boldsymbol{\zeta}) = \langle\Psi\rangle_{\boldsymbol{J}},$$

где $\mathrm{I}(\boldsymbol{\zeta}) \in \mathbb{R}^{m\times m}$ *– информационная матрица Фишера параметра* $\boldsymbol{\zeta}$ *модели* (7)*,* (8).

Доказательство. Информационная матрица Фишера $\mathrm{I}(\boldsymbol{\zeta})$ получается из информационной матрицы Фишера $\boldsymbol{J}$ параметра $\Gamma = \Psi\zeta$: $\mathrm{I}(\boldsymbol{\zeta}) = \Psi^{\dagger}\boldsymbol{J}\Psi = \langle\Psi\rangle_{\boldsymbol{J}}$. Тождество $\Sigma_{\boldsymbol{J}} = \mathrm{I}^{-1}(\boldsymbol{\zeta})$ получается прямой подстановкой $H = \boldsymbol{J}$ в (25). Следствие 3 доказано.

Пусть $M \subset \boldsymbol{M}_T^+ \subset \boldsymbol{X}^s$ – некоторое подмножество s-предысторий. Оценку (21) будем называть M-разреженной, если она использует только оценки (16) для $q \in M$. Это равносильно тому, что весовая матрица H в (20), (21) лежит в следующем множестве M-разреженных весовых матриц:

$$\boldsymbol{H}_M ::= \left\{H = H^{\dagger} \geq 0 : H_{q,q'} \equiv 0_{d\times d}, \forall (q,q') \notin M^2\right\}.$$

По своему построению оценка (21) $\boldsymbol{M}_T^+$-разрежена. Блочно-диагональная оптимальная весовая матрица (26), согласно (24), имеет все строго положительно определенные диагональные блоки $\boldsymbol{J}_{q,q} > 0,\ q \in \boldsymbol{X}^s$, поэтому она лежит в классе допустимых весовых матриц $\boldsymbol{H}_{\boldsymbol{M}_T^+}$ только при условии $\boldsymbol{M}_T^+ = \boldsymbol{X}^s$, которое на практике, как правило, достигается при очень больших значениях T длительности наблюдения временного ряда. Следующий результат доставляет оптимальную M-разреженную весовую матрицу.

Теорема 4. *Блочно-диагональная М-разреженная весовая матрица*

$$H = \boldsymbol{J}(M) \tag{27}$$

оптимальна в следующем смысле: для любой другой М-разреженной весовой матрицы $\tilde{H} \in \boldsymbol{H}_M$ *асимптотическая матрица ковариаций* (25) *М-разреженной оценки* (21), *отвечающей матрице* $\tilde{H}$, *удовлетворяет неравенству*

$$\Sigma_{\tilde{H}} \geq \Sigma_{\boldsymbol{J}(M)}.$$

Доказательство идентично доказательству теоремы 5 в [9] и опирается на лемму 5 из той же статьи.

Таким образом, оптимальная $\boldsymbol{M}_T^+$-разреженная весовая матрица, согласно (22), (24), (27), зависит от неизвестного значения параметра Θ и может быть статистически оценена подстановочным методом – заменами $\pi_q \to \hat{\pi}_q$, $\boldsymbol{\theta}^q \to \hat{\boldsymbol{\theta}}^q$ в формулах (22), (24) на основе оценок (15), (16).

Заключение. В работе получены следующие новые результаты:

1) для дискретных временных рядов предложена новая малопараметрическая модель DCNAR(*s,d*) на основе экспоненциального семейства дискретных распределений вероятностей с многомерным параметром;

2) построено семейство состоятельных асимптотически нормальных статистических оценок FBE явного вида для параметров предложенной модели ДВР;

3) в семействе FBE найдена асимптотически эффективная оценка (асимптотически достигающая границы Крамера – Рао).

Эти результаты могут быть использованы для робастного статистического анализа ДВР [9, 18, 19], статистического анализа дискретных пространственно-временных данных и случайных полей [20].

Благодарности. Работа выполнена в рамках проекта № 20211983 Государственной программы фундаментальных исследований «Конвергенция-2025».

Acknowledgements. This work was carried out within the project no. 20211983 of the State Program for Fundamental Research “Convergence-2025”.

Список использованных источников

1. Statistical analysis of multivariate discrete-valued time series / K. Fokianos [et al.] // J. Multivar. Anal. – 2022. – Vol. 188. – P. 104805. https://doi.org/10.1016/j.jmva.2021.104805

2. Kharin, Yu. Statistical analysis of discrete-valued time series by parsimonious high-order Markov chains / Yu. Kharin // Austr J. Statistics. – 2020. – Vol. 49, № 4. – P. 76–88. https://doi.org/10.17713/ajs.v49i4.1132

3. Anderson, T. W. The Statistical Analysis of Time Series / T. W. Anderson. – New York: Wiley, 1971. https://doi.org/10.1002/9781118186428

4. Box, G. Time Series Analysis: Forecasting and Control / G. Box, G. Jenkins. – San Francisco: Holden-Day, 1970.

5. Nelder, J. Generalized linear models / J. Nelder, R. Wedderburn // J. R. Stat. Soc., Ser. A. – 1972. – Vol. 135, № 3. – P. 370–384. https://doi.org/10.2307/2344614

6. Kedem, B. Regression Models for Time Series Analysis / B. Kedem, K. Fokianos. – Hoboken: Wiley, 2002. https://doi.org/10.1002/0471266981

7. Amari, S. Methods of Information Geometry / S. Amari, H. Nagaoka. – Oxford University Press, 2000. https://doi.org/10.1090/mmono/191

8. Yee, T. W. Vector Generalized Linear and Additive Models: With an Implementation in R / T. W. Yee. – New York: Springer, 2015. https://doi.org/10.1007/978-1-4939-2818-7

9. Kharin, Yu. Robust estimation for binomial conditionally nonlinear autoregressive time series based on multivariate conditional frequencies / Yu. Kharin, V. Voloshko // J. Multivar. Anal. – 2021. – Vol. 185. – P. 104777. https://doi.org/10.1016/j.jmva.2021.104777

10. Харин, Ю. С. Семибиномиальные условно нелинейные авторегрессионные модели дискретных случайных последовательностей: вероятностные свойства и статистическое оценивание параметров / Ю. С. Харин, В. А. Волошко // Дискрет. математика. – 2019. – Вып. 31, № 1. – С. 72–98.

11. Харин, Ю. С. Биномиальная условно нелинейная авторегрессионная модель дискретного временного ряда и ее вероятностно-статистические свойства / Ю. С. Харин, В. А. Волошко // Тр. Ин-та математики. – 2019. – Вып. 26, № 1. – С. 95–105.

12. Kharin, Yu. S. Statistical estimation of parameters for binary conditionally nonlinear autoregressive time series / Yu. S. Kharin, V. A. Voloshko, E. A. Medved // Math. Methods Stat. – 2018. – Vol. 27, № 2. – P. 103–118. https://doi.org/10.3103/s1066530718020023

13. Andersen, E. B. Sufficiency and exponential families for discrete sample spaces / E. B. Andersen // J. Am. Stat. Assoc. – 1970. – Vol. 65, № 331. – P. 1248–1255. https://doi.org/10.1080/01621459.1970.10481160

14. Kharin, Yu. S. Statistical analysis of conditionally binomial nonlinear regression time series with discrete regressors / Yu. S. Kharin, V. A. Voloshko // Theory Probab. Math. Stat. – 2020. – Vol. 100. – P. 181–190. https://doi.org/10.1090/tpms/1105

15. Engel, J. Polytomous logistic regression / J. Engel // Stat. Neerlandica. – 1988. – Vol. 42, № 4. – P. 233–252. https://doi.org/10.1111/j.1467-9574.1988.tb01238.x

16. Krisztin, T. A spatial multinomial logit model for analysing urban expansion / T. Krisztin, P. Piribauer, M. Wögerer // Spat. Econ. Anal. – 2022. – Vol. 17, № 2. – P. 223–244. https://doi.org/10.1080/17421772.2021.1933579

17. Billingsley, P. Statistical methods in Markov chains / P. Billingsley // Ann. Math. Stat. – 1961. – Vol. 32, № 1. – P. 12–40. https://doi.org/10.1214/aoms/1177705136

18. Kharin, Yu. Robustness in statistical pattern recognition under “contaminations” of training samples / Yu. Kharin, E. Zhuk // Proc. of the 12th IAPR Int. Conf. on Pattern Recognition. – 1994. – Vol. 2. – P. 504–506. https://doi.org/10.1109/icpr.1994.576996

19. Maevskii, V. V. Robust regressive forecasting under functional distortions in a model / V. V. Maevskii, Yu. S. Kharin // Automation and Remote Control. – 2002. – Vol. 63, № 11. – P. 1803–1820. https://doi.org/10.1023/a:1020959432568

20. Kharin, Yu. Statistical analysis of spatio-temporal data based on Poisson conditional autoregressive model / Yu. Kharin, M. Zhurak // Informatica. – 2015. – Vol. 26, № 1. – P. 67–87. https://doi.org/10.15388/informatica.2015.39

References

1. Fokianos K., Fried R., Kharin Yu., Voloshko V. Statistical analysis of multivariate discrete-valued time series. *Journal of Multivariate Analysis*, 2022, vol. 188, pp. 104805. https://doi.org/10.1016/j.jmva.2021.104805

2. Kharin Yu. Statistical analysis of discrete-valued time series by parsimonious high-order Markov chains. *Austrian Journal of Statistics*, 2020, vol. 49, no. 4, pp. 76–88. https://doi.org/10.17713/ajs.v49i4.1132

3. Anderson T. W. *The Statistical Analysis of Time Series*. New York, Wiley, 1971. https://doi.org/10.1002/9781118186428

4. Box G., Jenkins G. *Time Series Analysis: Forecasting and Control*. San Francisco, Holden-Day, 1970.

5. Nelder J., Wedderburn R. Generalized linear models. *Journal of the Royal Statistical Society. Series A*, 1972, vol. 135, no. 3, pp. 370–384. https://doi.org/10.2307/2344614

6. Kedem B., Fokianos K. *Regression Models for Time Series Analysis*. Hoboken, Wiley, 2002. https://doi.org/10.1002/0471266981

7. Amari S., Nagaoka H. *Methods of Information Geometry*. Oxford University Press, 2000. https://doi.org/10.1090/mmono/191

8. Yee T. W. *Vector Generalized Linear and Additive Models: With an Implementation in R*. New York, Springer, 2015. https://doi.org/10.1007/978-1-4939-2818-7

9. Kharin Yu., Voloshko V. Robust estimation for binomial conditionally nonlinear autoregressive time series based on multivariate conditional frequencies. *Journal of Multivariate Analysis*, 2021, vol. 185, pp. 104777. https://doi.org/10.1016/j.jmva.2021.104777

10. Kharin Yu. S., Voloshko V. A. Semibinomial conditionally nonlinear autoregressive models of discrete random sequences: probabilistic properties and statistical parameter estimation. *Discrete Mathematics and Applications*, 2020, vol. 30, no. 6, pp. 417–437. https://doi.org/10.1515/dma-2020-0038

11. Kharin Yu., Voloshko V. Binomial conditionally nonlinear autoregressive model of discrete-valued time series and its probabilistic and statistical properties. *Trudy Instituta matematiki = Proceedings of the Institute of Mathematics*, 2019, vol. 26, no. 1, pp. 95–105 (in Russian).

12. Kharin Yu. S., Voloshko V. A., Medved E. A. Statistical estimation of parameters for binary conditionally nonlinear autoregressive time series. *Mathematical Methods of Statistics*, 2018, vol. 27, no. 2, pp. 103–118. https://doi.org/10.3103/s1066530718020023

13. Andersen E. B. Sufficiency and exponential families for discrete sample spaces. *Journal of the American Statistical Association*, 1970, vol. 65, no. 331, pp. 1248–1255. https://doi.org/10.1080/01621459.1970.10481160

14. Kharin Yu. S., Voloshko V. A. Statistical analysis of conditionally binomial nonlinear regression time series with discrete regressors. *Theory of Probability and Mathematical Statistics*, 2020, vol. 100, pp. 181–190. https://doi.org/10.1090/tpms/1105

15. Engel J. Polytomous logistic regression. *Statistica Neerlandica*, 1988, vol. 42, no. 4, pp. 233–252. https://doi.org/10.1111/j.1467-9574.1988.tb01238.x

16. Krisztin T., Piribauer P., Wögerer M. A spatial multinomial logit model for analysing urban expansion. *Spatial Economic Analysis*, 2022, vol. 17, no. 2, pp. 223–244. https://doi.org/10.1080/17421772.2021.1933579

17. Billingsley P. Statistical methods in Markov chains. *The Annals of Mathematical Statistics*, 1961, vol. 32, no. 1, pp. 12–40. https://doi.org/10.1214/aoms/1177705136

18. Kharin Yu., Zhuk E. Robustness in statistical pattern recognition under "contaminations" of training samples. *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, 1994, vol. 2, pp. 504–506. https://doi.org/10.1109/icpr.1994.576996

19. Maevskii V. V., Kharin Yu. S. Robust regressive forecasting under functional distortions in a model. *Automation and Remote Control*, 2002, vol. 63, no. 11, pp. 1803–1820. https://doi.org/10.1023/a:1020959432568

20. Kharin Yu., Zhurak M. Statistical analysis of spatio-temporal data based on Poisson conditional autoregressive model. *Informatica*, 2015, vol. 26, no. 1, pp. 67–87. https://doi.org/10.15388/informatica.2015.39

Информация об авторах

Волошко Валерий Анатольевич – кандидат физико-математических наук, заведующий сектором компьютерного анализа данных, Научно-исследовательский институт прикладных проблем математики и информатики, Белорусский государственный университет (пр. Независимости, 4, 220030, Минск, Республика Беларусь). E-mail: valeravoloshko@yandex.ru

Харин Юрий Семенович – академик Национальной академии наук Беларуси, доктор физико-математических наук, профессор, директор Научно-исследовательского института прикладных проблем математики и информатики, Белорусский государственный университет (пр. Независимости, 4, 220030, Минск, Республика Беларусь). E-mail: kharin@bsu.by

Information about the authors

Valeriy A. Voloshko – Ph. D. (Physics and Mathematics), Head of the Computer Data Analysis Sector, Research Institute for Applied Problems of Mathematics and Informatics, Belarusian State University (Nezavisimosty Ave., 4, Minsk, 220030, Republic of Belarus).
E-mail: valeravoloshko@yandex.ru

Yuriy S. Kharin – Academician of the National Academy of Sciences of Belarus, Dr. Sc. (Physics and Mathematics), Professor, Head of the Research Institute for Applied Problems of Mathematics and Informatics, Belarusian State University (Nezavisimosty Ave., 4, Minsk, 220030, Republic of Belarus). E-mail: kharin@bsu.by