КОМПЬЮТЕРНЫЕ ИССЛЕДОВАНИЯ
И МОДЕЛИРОВАНИЕ 2026 Т. 18 № 4 С. 1021–1034
DOI: 10.20537/2076-7633-2026-18-4-1021-1034

Ки&М

МОДЕЛИ ЭКОНОМИЧЕСКИХ И СОЦИАЛЬНЫХ СИСТЕМ

УДК: 51-77; 330.3

Идентификация типа распределения и его применение при оценке функционирования сложных систем на основе байесовских интеллектуальных технологий

Р. А. Жуков[1,2,a], А. Д. Прачев[2], Е. А. Ламзина[2]

[1]Финансовый университет при Правительстве Российской Федерации (Тульский филиал),
Россия, 300012, г. Тула, ул. Оружейная, д. 1а
[2]Финансовый университет при Правительстве Российской Федерации,
Россия, 125167, г. Москва, пр-кт Ленинградский, д. 49/2

E-mail: [a] pluszh@mail.ru

Получено 21.04.2026, после доработки — 17.07.2026.
Принято к публикации 21.07.2026.

В настоящей работе представлена методика оценки функционирования сложных систем с учетом типа распределения показателей, характеризующих функционирование структурных элементов такой системы, действующей в условиях неопределенности. В отличие от ранее проведенных исследований, в работе осуществлена интеграция идентификации типа распределения индикаторов с их байесовской оценкой с учетом нормативов, характеризующих требуемые результаты функционирования структурных элементов системы. Тип распределения и его характеристики дают более полную информацию о свойствах объекта, и их учет влияет на результаты оценки функционирования элементов сложных систем, особенно в случаях наличия «тяжелых» хвостов, характерных для ряда социально-экономических процессов и систем. Показано, что идентификацию типа распределения целесообразно осуществлять по модифицированным за счет исключения наилучшего по статистическим характеристикам тренда данным. Определение наилучшего типа распределения осуществляется по критерию согласия Колмогорова – Смирнова, показавшему лучшие (по имеющееся выборке), по сравнению с другими критериями или их смесью, статистические оценки. Полученный тип распределения используется при вероятностной оценке результатов функционирования элементов подсистемы на основе байесовских интеллектуальных технологий, успешно работающих в условиях неопределенности и которые дают возможность представить такие результаты на числовой и лингвистической шкалах в общей иерархической информационной модели. Методика была апробирована на примере объектной подсистемы (пространственно-временная классификация Г. Б. Клейнера) экономической подсистемы Тульской области. В качестве показателей оценки были использованы объемы валового регионального продукта по шести разделам общероссийского классификатора видов экономической деятельности (разделы A, B, C, D, E). В некоторых случаях удается снизить уровень неопределенности в оценках показателей; получить более строгие по отношению к нормативу оценки, что в итоге позволяет сделать вывод, что использование типа распределения при формировании вероятностных оценок показателей структурных элементов сложных систем дает возможность обосновать их корректное применение в рамках концепции доказательного моделирования.

Ключевые слова: тип распределения, выборка, тренд, модель, байесовские интеллектуальные технологии

Статья подготовлена по результатам исследований, выполненных за счет бюджетных средств по государственному заданию Финуниверситета.

COMPUTER RESEARCH AND MODELING
2026 VOL. 18 NO. 4 P. 1021–1034
DOI: 10.20537/2076-7633-2026-18-4-1021-1034

Ки&М

MODELS OF ECONOMIC AND SOCIAL SYSTEMS

UDC: 51-77; 330.3

Identification of the type of distribution and its application in assessing the functioning of complex systems based on Bayesian intelligent technologies

R. A. Zhukov[1,2,a], A. D. Prachev[2], E. A. Lamzina[2]

[1]Financial University under the Government of the Russian Federation (Tula Brunch),
1a Oruzheynaya st., Tula, 300012, Russia
[2]Financial University under the Government of the Russian Federation,
49/2 Leningradsky pr., Moscow, 125167, Russia

E-mail: [a] pluszh@mail.ru

Received 21.04.2026, after completion — 17.07.2026.
Accepted for publication 21.07.2026.

This study presents a methodology for assessing the functioning of complex systems, taking into account the type of distribution of indicators characterizing the functioning of structural elements of such a system operating under conditions of uncertainty. Unlike previous studies, the work integrates the identification of the type of indicator distribution with their Bayesian assessment, taking into account the norms characterizing the required results of the functioning of the structural elements of the system. The type of distribution and its characteristics provide more complete information about the properties of an object and their consideration affects the results of evaluating the functioning of elements of complex systems, especially in cases of the presence of "heavy" tails characteristic of a number of socio-economic processes and systems. It is shown that it is advisable to identify the type of distribution based on modified data due to the exclusion of the best trend in statistical characteristics. The determination of the best type of distribution is carried out according to the Kolmogorov – Smirnov agreement criterion, which showed the best (according to the available sample) statistical estimates compared to other criteria or a mixture of them. The obtained type of distribution is used in the probabilistic assessment of the results of the functioning of subsystem elements based on Bayesian intelligent technologies that successfully operate in conditions of uncertainty, and which make it possible to present such results on numerical and linguistic scales in a general hierarchical information model. The methodology was tested using the example of the object subsystem (G. B. Kleiner's spatio-temporal classification) of the economic subsystem of the Tula region, considered as a complex system — a socio-ecological-economic system. The volume of gross domestic product by region in six sections of the all-Russian classifier of economic activities (sections A, B, C, D, E) was used as assessment indicators. In some cases, it is possible to reduce the level of uncertainty in the estimates of indicators.; to obtain estimates that are more stringent than the norm, which ultimately leads to the conclusion that the use of the distribution type in the formation of probabilistic estimates of indicators of structural elements of complex systems makes it possible to justify their correct application within the framework of the concept of evidence-based modeling.

Keywords: distribution type, sampling, trend, model, Bayesian intelligent technologies

Citation: *Computer Research and Modeling*, 2026, vol. 18, no. 4, pp. 1021–1034 (Russian).

The article was prepared based on the results of research carried out at the expense of budgetary funds on the state assignment of the Financial University.

Введение

Идентификация типа распределения (в том числе его характеристик) показателей социально-экономических и других процессов сложных систем является одним из возможных ключевых элементов методологии научного познания, позволяющего понять природу явлений, с целью формирования базы для обоснования корректности проводимых исследований и интерпретации полученных результатов, в том числе в рамках концепции доказательного моделирования [Клейнер, 2023]. В работе [Taleb, 2025] представлены классы возможных статистических распределений с акцентом на распределения с «тяжелыми» хвостами, что характерно для ряда социально-экономических систем и процессов. Разработанный инструментарий используется для формирования статистических выводов и принятия решений, однако не применяется в сочетании с нормативами (заданными, плановыми, значениями), позволяющими качественно оценить результаты функционирования структурных элементов системы, в том числе с использованием лингвистических шкал.

В большинстве исследований акцент ставится либо на идентификацию типа распределения [Silbersdorff, Schneider, 2019; Charpentier, Flachaire, 2022; Mirzaei, Jahanshahi, 2022], либо на использование априорно заданных типов (в том числе смеси распределений) и их применение при моделировании сложных систем [Iacopini et al., 2023; Angelini et al., 2024; Ramos et al., 2024]. Моделирование с использованием различных типов распределений на основе вариационного байесовского алгоритма с целью получения адекватных регрессионных моделей для прогнозирования инфляции отражено в [Koop, Korobilis, 2023]. При этом не ставится вопрос о качественной оценке (например, выше или ниже норматива) показателя.

Базовой проблемой идентификации типа распределения является характер выборки (секционная, кросс-секционная, панельные данные, временной ряд), которую целесообразно применять в проводимых измерениях, что требует введения некоторых допущений.

Еще одной проблемой в выборках показателей, характеризующих региональный уровень социально-экономического развития территорий, является ограниченность данных. В этом случае для идентификации типа распределения можно воспользоваться соответствующей методикой, которая определяет тип распределения по асимметрии и эксцессу выборочных данных, размещенных на плоскости Пирсона [Прокопчина, 2020]. Однако в некоторых случаях по имеющемуся набору данных не удается идентифицировать распределение в связи с ограниченным количеством классов (25 классов), для которых рассчитаны разделяющие границы эталонных классов [Жуков и др., 2022].

Кроме того, статистические данные, в частности данные о функционировании субъектов Российской Федерации, обладают в некотором смысле неопределенностью, что приводит к необходимости оценивать результаты функционирования субъектов Российской Федерации, рассматриваемых как сложные системы (в том числе их структурные элементы) — социо-эколого-экономические системы (СЭЭС), — в терминах теории вероятностей и математической статистики [Айвазян, Мхитарян, 1998] и применять байесовские интеллектуальные технологии (далее — БИТ) как инструмент корректного и обоснованного анализа и интерпретации результатов [Прокопчина, 2021а]. В этом аспекте включение этапа идентификации типов распределений и их характеристик (в некоторых исследованиях тип распределения задается априори) в общий алгоритм оценки состояния и функционирования систем на основе БИТ является важным элементом методологии в рамках концепции доказательного моделирования.

Таким образом, целью исследования является реализация БИТ с включением этапа идентификации типа распределения показателей, характеризующих состояние и функционирование СЭЭС.

Методология и информационная база исследования

Методология исследования включает в себя следующие этапы.

1. Определение объекта исследования. В качестве объекта исследования выступает сложная система (например, субъект Российской Федерации), включающая в себя взаимосвязанные подсистемы и элементы, входящие в состав подсистем. Каждый из i элементов ($i = 1, \ldots, I \in N$) характеризуется признаковым описанием, или признаком y_i (будем рассматривать его как случайную величину). Значение признака $y_{i,t}$ характеризует результат функционирования элемента i в период времени t ($t = 1, \ldots, T \in N$). Таким образом, объект исследования определяется набором признаковых описаний, в данном случае — результативных признаков y_i.

2. Определение типа выборки. Обоснование типа выборки основано на следующих предпосылках и допущениях. Формирование выборки определяется целями исследования, особенностями объекта и предмета исследования, что в свою очередь определяет структуру данных. Будем предполагать, что каждый субъект Российской Федерации обладает своими специфическими (уникальными) условиями функционирования. Следовательно, каждый субъект должен иметь свои собственные (в зависимости от условий функционирования) нормативы, позволяющие классифицировать уровень функционирования, его подсистем и элементов $\overline{y}_{i,t}$ (от предельно ниже нормы до предельно выше нормы [Прокопчина, 2021b]) [Жуков, Прокопчина, 2024a]. Следует отметить, что под нормой в БИТ понимается заранее установленное, заданное (плановое) значение или норматив, при котором функционирование структурного элемента системы считается удовлетворительным. Тогда оценка функционирования структурных элементов системы должна осуществляться по данным, соответствующим этой системе. При этом данные с большой вероятностью содержат тренды, что приводит к необходимости их исключения при идентификации типа распределения результативных признаков. Однако для показателей регионального уровня, размещенных в открытых источниках, обычно используют годовые значения, что приводит к использованию «коротких» временных рядов при построении трендовых моделей, но с достаточным числом наблюдений, чтобы построить такие модели и оценить их характеристики.

3. Идентификация тренда с учетом типа распределения. Для построения тренда можно воспользоваться традиционными методами, в частности методом наименьших квадратов (далее — МНК, OLS) или методом максимального правдоподобия (далее — ММП, MLE), последний из которых не накладывает ограничения на тип распределения случайной величины.

Поскольку тип распределения изначально неизвестен, функцию максимального правдоподобия можно построить для различных типов распределения, из которых можно выбрать наилучшую модель с учетом типа распределения. В простейшем случае можно воспользоваться информационным критерием Акаике [Akaike, 1974]:

$$AIC_{i,j^*,m^*} = \min_{j,m}(AIC_{i,j,m}) = \min_{j,m}(2 \cdot k_{i,j,m} - 2 \cdot \ln(L_{i,j,m})), \tag{1}$$

где $k_{i,j,m}$ — число параметров модели тренда, $L_{i,j,m}$ — максимизированное значение функции правдоподобия модели, j — тип распределения случайной величины ($j = 1, \ldots, J \in N$); m — функциональная форма тренда ($m = 1, \ldots, M \in N$), i — номер элемента (признака), m^* — выбранная функциональная форма модели тренда, j^* — выбранный тип распределения.

При этом $L_{i,j,m}$ определяется по формуле

$$L_{i,j,m} = \max_{\theta_m \in \Theta_m} \left[\prod_{t=1}^{T} p_j(y_{i,t}, \theta_{i,j,m}) \right], \tag{2}$$

где $y_{i,t}$ — значение признака i в период времени t; $\theta_{i,j,m}$ — вектор параметров модели тренда m; Θ_m — множество значений параметров θ_m; p_j — плотность распределения вероятностей типа j.

Тогда оценка $\widehat{\theta}_{i,j,m}$ вектора параметров $\theta_{i,j,m}$ модели будет определяться по формуле

$$\widehat{\theta}_{i,j,m} = \underset{\theta_{i,j,m} \in \Theta_{i,j,m}}{\arg\max} [L_{i,j,m}]. \tag{3}$$

В случае малых выборок можно использовать критерий Акаике с поправкой на их объем:

$$AICc_{i,j^*,m^*} = \min_{j,m} \left(AIC_{i,j,m} + \frac{2 \cdot k_{i,j,m}^2 + 2 \cdot k_{i,j,m}}{n - k_{i,j,m} - 1} \right), \tag{4}$$

где n — мощность выборки.

4. Исключение тренда из данных. По сути, тренд представляет собой неслучайную величину, зависящую от параметра времени t. Тогда будем считать, что истинным типом распределения случайной величины y_i будет являться тип распределения, соответствующий случайной величине ε_i, значения которой определяются по формуле

$$\varepsilon_{i,t} = y_{i,t} - f_{i,j^*,m^*} \left(\widehat{\theta}_{i,j^*,m^*}, t \right), \tag{5}$$

где $f_{i,j^*,m^*} \left(\widehat{\theta}_{i,j^*,m^*}, t \right)$ — трендовая модель с функциональной формой m^* с учетом типа распределения j^*.

5. Идентификация типа распределения случайной величины. Для определения типа распределения случайной величины будем использовать критерий Колмогорова – Смирнова [Лемешко, 2024], который рассчитывается по формуле

$$D_{i,j,T} = \sup_t |F_T(\varepsilon_{i,t}) - F_m(\varepsilon_{i,t})|, \tag{6}$$

где $F_T(\varepsilon_{i,t})$ — значение эмпирической функции распределения случайной величины ε_i в период времени t (наблюдение t), $F_m(\varepsilon_{i,t})$ — значение теоретической функции распределения случайной величины ε_i в период времени t (наблюдение t) с типом m.

Значимость $p - value_{i,m}$ (нулевая гипотеза: соответствие эмпирического распределения теоретическому распределению) определяется по формуле

$$p - value_{i,m} = 1 - K \cdot \exp\left(-2 \cdot D_{i,m,T}^2 \cdot T \right), \tag{7}$$

где K — корректирующий коэффициент, T — число наблюдений.

Тогда наилучший тип распределения m^{**} может быть определен по максимальному уровню значимости по формуле

$$p - value_{i,m^{**}} = \max(p - value_{i,1}, \ldots, p - value_{i,m}). \tag{8}$$

В случае если $p - value_{i,m^{**}}$ не превышает порога значимости, например 0,05, то в качестве альтернативы можно использовать критерий AIC.

Замечание. Также можно использовать другие критерии (Андерсона – Дарлинга — более чувствителен к отклонениям в хвостах распределения; Крамера – фон Мизеса — учитывает различия по всей области определения; χ^2 — классический критерий и ряд других) или их смесь. Однако при тестировании применения смеси критериев на модельных данных оказалось, что смесь уменьшает процент совпадения определенного по критериям типа распределения сгенерированных данных того же типа распределения. При этом использование критерия Колмогорова – Смирнова давало больший процент совпадения по сравнению с другими критериями. Модельный эксперимент проводился на сгенерированных (использован встроенный алгоритм генерации в Python) выборках 10, 50, 100 наблюдений для каждого из 9 распределений с 5 повторениями. Процент совпадений заданного и оцененного распределений при использовании

критерия Колмогорова – Смирнова составил порядка 75–80 %, при использовании других критериев или их смеси процент совпадений — 60–70 %. Следует отметить, что такой результат может быть также связан и с особенностями встроенного в Python алгоритма выборки с заданным типом распределения, вносящего возможные искажения в результат оценки. В дальнейшем предполагается провести модельный эксперимент на данных, полученных с помощью других алгоритмов и инструментальных средств, например R.

6. Представление значений признаков в вероятностном виде [Прокопчина, 2021a; Жуков, Прокопчина, 2024a]. Пусть имеется набор i ($i = 1, \ldots, I \in N$) признаков y_i со значениями $y_{i,t}$ в период времени t ($t = 1, \ldots, T \in N$). При этом $y_{i,t} \in [y_{i,\min} - \sigma_i; y_{i,\max} + \sigma_i]$, σ_i — расширение (корректировка) диапазона числовой шкалы, имеющего смысл динамического ограничения для y_i. Пусть $\forall y_{i,t}$ имеется нормативное значение $y_{i,\text{norm}}$, которое известно (задается априорно) или в простейшем случае определяется по формуле

$$y_{i,\text{norm}} = \frac{y_{i,\min} + y_{i,\max}}{2}. \tag{9}$$

Корректировка диапазона σ_i определяется по формуле

$$\sigma_i = \frac{y_{i,\text{norm}}}{3}. \tag{10}$$

В более сложном случае норматив может быть установлен с учетом влияния факторов (условий) на признак y_i, в том числе с учетом методик, основанных на построении регрессионных моделей [Журавлев, Жуков, 2012; Жуков, 2025b], или методики, учитывающей корреляционные зависимости результативного признака и факторов [Жуков, 2025a].

В рамках БИТ будем представлять значения признака $y_{i,t}$ в виде набора пар значений $\left(h_{i,t,j_p}; p_{i,t,j_p}\right)$, где h_{i,t,j_p} — значение признака $y_{i,t}$, соответствующее положению репера на числовой шкале и классу на сопряженной ей лингвистической шкале (представляет собой набор 9 классов от предельно ниже нормы до предельно выше нормы — выше/ниже нормативного значения, — из множества решений — альтернативных (эталонных) гипотез H_{i,t,j_p} ($j_p = 1, \ldots, J_p \in N$); p_{i,t,j_p} — вероятность того, что признак $y_{i,t}$ примет значение h_{i,t,j_p}. То есть гипотез таких, что $H_{i,t,j_p}: y_{i,t} = h_{i,t,j_p}$ с эталонным распределением, соответствующим типу распределения m^{**}.

Априорная вероятность (a — обозначение априорной вероятности) эталонной гипотезы определяется по формуле

$$p^a_{i,t,j_p}\left(H_{i,t,j_p}: y_{i,t} = h_{i,t,j_p}\right) \equiv p^a_{i,t,j_p}\left(H_{i,t,j_p}\right) = p^a_{i,t,j_p} = \frac{1}{J_p}. \tag{11}$$

Тогда апостериорная вероятность (ap — обозначение апостериорной вероятности), то есть вероятность при условии, что произошло событие S, которое заключается в совместном появлении оценок $\widetilde{h}_{i,t,j_p}$, будет определяться по формуле Байеса:

$$p^{ap}_{i,t,j_p}\left(H_{i,t,j_p}|S\right) = \frac{p^a_{i,t,j_p}\left(H_{i,t,j_p}\right) \cdot p\left(S|H_{i,t,j_p}\right)}{\sum\limits_{j_p=1}^{J_p} p^a_{i,t,j_p}\left(H_{i,t,j_p}\right) \cdot p\left(S|H_{i,t,j_p}\right)}, \tag{12}$$

где $p\left(S|H_{i,t,j_p}\right)$ — условная вероятность гипотезы при появлении события S, которая определяется по формуле

$$p\left(S|H_{i,t,j_p}\right) = p_{m^{**}}\left(y_{i,t}, h_{i,t,j_p}\right), \tag{13}$$

где $p_{m^{**}}\left(y_{i,t},\, h_{i,t,j_p}\right)$ — значение плотности вероятности наилучшего распределения типа m^{**} для значения $y_{i,t}$ и математического ожидания h_{i,t,j_p}.

В первом приближении среднеквадратическое отклонение, которое соответствует шагу между реперами, может быть определено по формуле

$$\sigma_{m^{**}}(y_{i,t}) = \frac{y_{i,\max} - y_{i,\min}}{J_p}. \tag{14}$$

Такой выбор обусловлен тем, что число реперов (как и количество классов) обычно выбирается равным 9, что в сочетании с диапазоном числовой шкалы и формулами (8) и (9) дает отклонение $y_{i,\min}$ и $y_{i,\max}$ от среднего значения, в данном случае от $y_{i,\text{norm}}$ признака y_i порядка $\pm 3 \cdot \sigma_i$ — хорошо работающее правило для нормального распределения. Такое определение $\sigma_{m^{**}}(y_{i,t})$ можно заменить на расчет стандартного отклонения непосредственно по значениям $y_{i,t}$, так же как и другие характеристики типа распределения m^{**}.

Таким образом, в рамках представленной методологии БИТ вероятностные оценки признака со значениями $y_{i,t}$ определяются с учетом типа распределения $y_{i,t}$.

Совокупность признаков, отражающих поведение элементов системы и которые могут быть сгруппированы в подсистемы и систему в целом, можно представить в виде иерархической информационной модели с визуализацией результатов функционирования, как показано, например, в [Жуков, 2025a].

В качестве инструментальных средств были использованы: программная платформа «Инфоаналитик 2.0» (для построения вероятностных моделей и визуализации результатов оценки) [Жуков, Прокопчина, 2024b] с дополнительным модулем автоматического подбора типа распределения признаков, разработанным на языке программирования Python в среде PyCharm [PyCharm IDE, 2026].

В качестве информационной базы были использованы данные из открытых источников за 2017–2023 гг. [Регионы России, 2026; Инфляция в России, 2026].

Результаты и обсуждение

В рамках исследования в качестве результативных признаков были выбраны объемы валового регионального продукта (далее — ВРП) по видам экономической деятельности ОКВЭД 2, характеризующие объектную подсистему пространственно-временной классификации [Клейнер, Рыбачук, 2017; Клейнер, Рыбачук, 2019], которая включает сельское хозяйство (раздел A), добычу полезных ископаемых (раздел B), обрабатывающие производства (раздел C), обеспечение электроэнергией, газом и паром (раздел D), водоснабжение, водоотведение и обращение с отходами (раздел E). Значения признаков были скорректированы на уровень инфляции и приведены к уровню 2017 г. Нормативы для каждого из показателей установлены по 2022 году и рассчитывались с учетом влияющих факторов [Жуков, 2025a]. Были выбраны 2 функциональные формы тренда (линейная и квадратичная) и использованы 9 типов распределения: нормальное (Н), логнормальное (Л), экспоненциальное (Э), гамма (Г), бета (Б), Вейбулла (В), треугольное (Т), Симпсона (С) и равномерное (Р). Результаты оценки наилучших трендов и идентификации лучшего типа распределения на примере Тульской области приведены в таблице 1. Для каждого из разделов вторая строка соответствует результатам при использовании критерия Акаике с поправкой на малый объем выборки. В случае идентичности результатов соответствующие ячейки объединены в одну ячейку.

Из таблицы 1 видно, что использование модифицированного критерия с учетом объема выборки не приводит к изменению суждения о функциональной форме тренда и типе распределения остатков. В данном исследовании значение критерия для остатков меняется с точностью до

Таблица 1. Результаты оценки наилучших трендов и идентификация наилучшего типа распределения на примере Тульской области по данным за 2017–2023 гг.

П	ТТ	ТР	R^2	AIC-T	TPO	p	AIC-O	TPA	pA	AIC-A
A	Л	Б	0,411	63,961	Л	0,817	148,245	Н	0,646	144,832
		Р		162,606			156,245			147,831
B	К	Б	0,953	75,929	Э	0,967	98,848	Н	0,928	101,689
				−36,071			101,848			104,689
C	К	Б	0,907	150,936	Г	0,996	158,466	Н	0,944	156,666
				38,936			166,466			159,666
D	К	Р	0,924	113,812	В	0,998	112,372	Н	0,970	111,623
		Б		25,358			120,372			114,623
E	Л	Э	0,323	105,726	Г	0,682	65,046	Н	0,475	110,626
				125,726			73,046			113,626

Примечание. П — показатель (буквенное обозначение — наименование раздела по ОКВЭД 2); ТТ — тип тренда (Л — линейный, К — квадратичный); ТР — тип распределения, использованный в методе MLE для оценки параметров тренда; R^2 — коэффициент детерминации; AIC-T — значение критерия Акаике для трендовой модели; TPO — тип распределения остатков; p — p-value; AIC-O — значение критерия Акаике при оценке типа распределения; TPA — альтернативный тип распределения; pA — p-value альтернативного типа распределения; AIC-A — значение критерия Акаике при оценке альтернативного типа распределения; Б — бета-распределение; В — распределение Вейбулла; Г — гамма-распределение; Л — логнормальное распределение; Н — нормальное распределение; Р — равномерное распределение; Э — экспоненциальное распределение.

константы. Поскольку число наблюдений невелико, то более предпочтительным является использование *AICc*, хотя принципиального значения в данном исследовании выбор *AICc* вместо *AIC* не имеет.

К сожалению, имеющиеся данные, в части малого числа наблюдений (всего 7 наблюдений), снижают надежность полученных оценок. Преодолением такого ограничения может служить применение идентификации типа распределения по высшим моментам с использованием плоскости Пирсона [Прокопчина, 2020]. Однако такое применение требует расширения количества эталонных классов типов распределений с разделяющими границами (в настоящее время представлено 25 классов), что требует проведения серьезного исследования. Вообще говоря, требуется не менее 7000 наблюдений для корректного применения методов математической статистики, как показано в [Прокопчина, 2020].

На рис. 1 представлены результаты оценки тренда и типа распределения остатков с демонстрацией автокорреляции (для различного количества лагов) для объема ВРП (раздел С. Обрабатывающие производства) Тульской области по данным за 2017–2023 гг. Из рис. 1 видно, что наилучшим типом распределения остатков оказалось распределение, которое, в соответствии с методологией, использовано для формирования вероятностной оценки функционирования элемента экономической подсистемы Тульской области, под которым понимается совокупность институциональных единиц, резидентов субъекта Российской Федерации, вносящих вклад в ВРП региона.

В дополнение к приведенным расчетам была проведена оценка ряда остатков на автокорреляцию и стационарность (разработанный модуль был расширен соответствующими тестами). Для проверки на автокорреляцию использованы тесты Дарбина – Уотсона и Льюнга – Бокса. Стационарность проверялась с помощью тестов Дики – Фуллера и KPSS, являющегося дополняющим к предыдущему тесту, поскольку имеет противоположные Дики – Фуллеру гипотезы. Результаты оценки представлены в таблице 2.

Как видно из таблицы 2, полученные результаты неоднозначны в части одновременного выполнения отсутствия автокорреляции и стационарности (например, по одному критерию ряд остатков стационарен, по другому критерию — нестационарен). Прежде всего это обусловлено

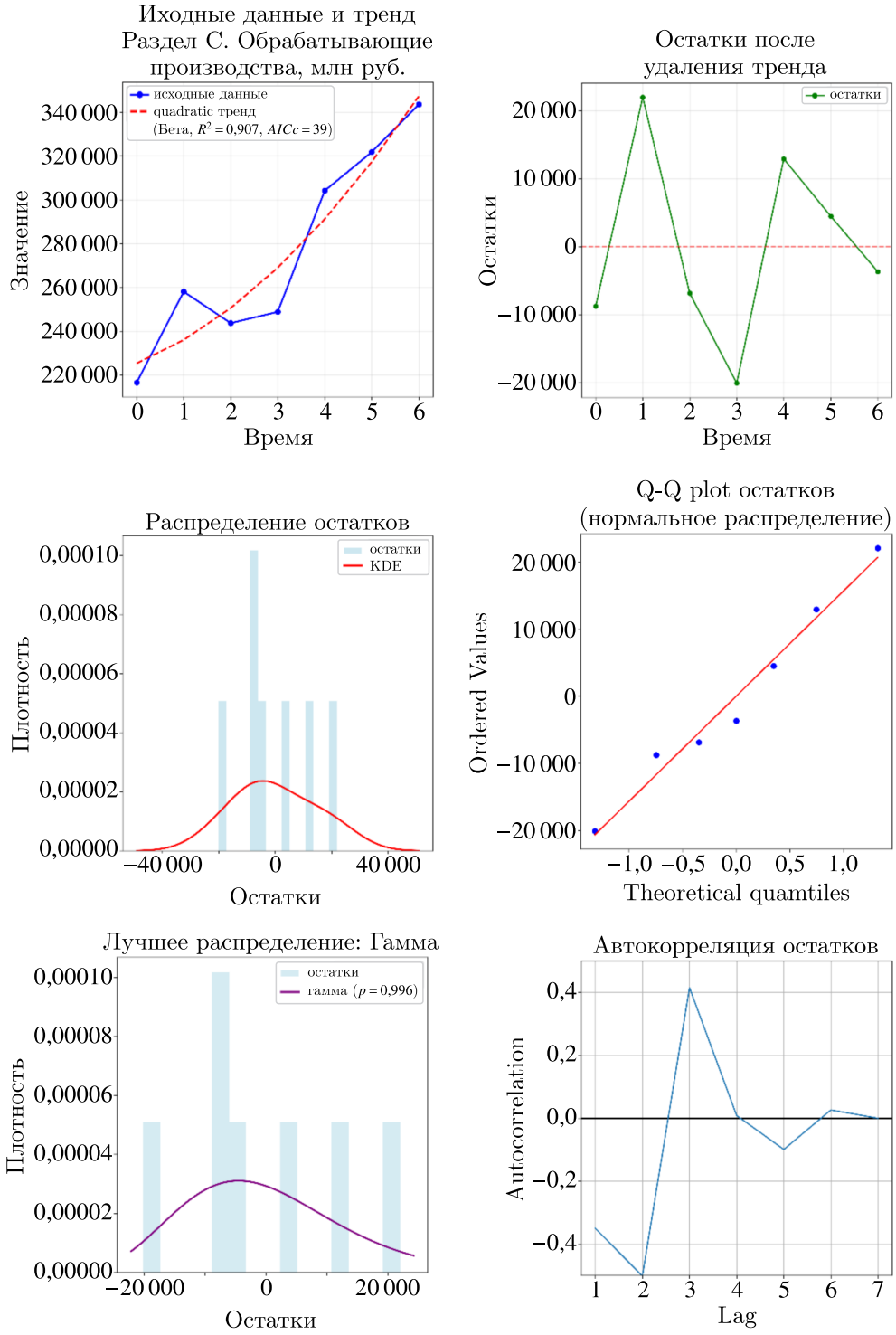

Рис. 1. Результаты оценки наилучших тренда и типа распределения остатков для раздела С Тульской области по данным за 2017–2023 гг.

априорным выбором функциональной формы тренда, а также ограниченным объемом выборки. Однако это не мешает общей демонстрации методики и принятию авторами рисков в части надежности полученных результатов. В дальнейшем рассмотрение других функциональных форм моделей может привести к улучшенным оценкам автокорреляции и стационарности.

Таблица 2. Результаты тестирования ряда остатков на автокорреляцию и стационарность на примере Тульской области по данным за 2017–2023 гг.

П	DW	A_{DW}	LB	p_{LB}	ADF	p_{ADF}	S_{ADF}	KPSS	p_{KPSS}	S_{KPSS}
A	0,942	нет	1,136	0,286	−1,915	0,325	нет	0,134	0,100	да
B	2,201	да	0,316	0,574	−2,248	0,190	нет	0,094	0,100	да
C	2,625	нет	1,285	0,257	−5,054	0,000	да	0,322	0,100	да
D	2,580	нет	1,905	0,168	−3,192	0,020	да	0,500	0,042	да
E	2,172	да	0,159	0,690	−2,052	0,264	нет	0,0968	0,100	да

Примечание: П — показатель (буквенное обозначение — наименование раздела по ОКВЭД 2); DW — значение критерия Дарбина – Уотсона; A_{DW} — автокорреляция отсутствует (да) / присутствует (нет) на уровне значимости 0,05; LB — значение критерия Льюинга – Бокса (для лага $k = 1$); p_{LB} — p-value для LB (автокорреляция отсутствует на уровне $p < p_{LB}$); ADF — статистика Дики – Фуллера; p_{ADF} — p-value для ADF (ряд стационарен на уровне значимости $p < p_{ADF}$); S_{ADF} — ряд стационарен на уровне значимости 0,05 (да/нет); KPSS — статистика теста KPSS; p_{KPSS} — p-value для KPSS (ряд стационарен на уровне значимости $p > p_{KPSS}$); S_{KPSS} — ряд стационарен на уровне значимости 0,05 (да/нет).

На рис. 2 и рис. 3 на лингвистической шкале представлены вероятностные оценки объема ВРП для раздела С. Обрабатывающие производства Тульской области в 2023 г. с использованием альтернативного (нормального) и наилучшего (гамма) типов распределения.

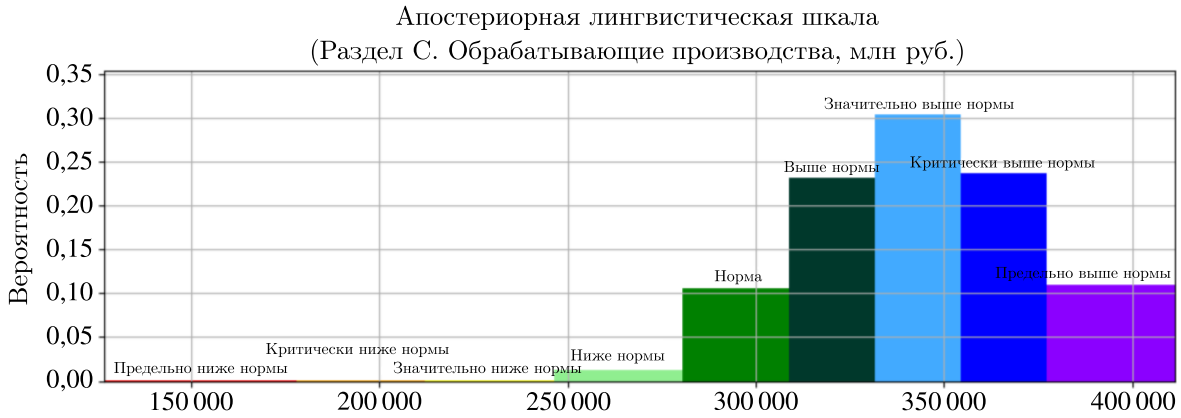

Рис. 2. Значения объема ВРП для раздела С. Обрабатывающие производства Тульской области в 2023 г. на лингвистической шкале при использовании альтернативного (нормального) распределения

Рис. 3. Значения объема ВРП для раздела С. Обрабатывающие производства Тульской области в 2023 г. на лингвистической шкале при использовании наилучшего (гамма) распределения

Во втором варианте наблюдаются меньшие значения вероятностей для классов, ближайших к классу с наибольшим значением вероятности, что уменьшает уровень неопределенности показателя. Действительно, в случае использования нормального распределения вероятности, превышающие заданный (установленный) порог значимости 0,15, обнаруживаются для классов «выше нормы», «значительно выше нормы» и «критически выше нормы». То есть результат функционирования элемента можно со значимой вероятностью оценить тройкой классов (или «выше нормы», или «значительно выше нормы», или «критически выше нормы»). В этом случае, формально уровень неопределенности можно оценить как $\frac{3}{9}$. При использовании гамма-распределения значимая вероятность характерна только для одного класса («значительно выше нормы»), а уровень неопределенности в таком контексте составит $\frac{1}{9}$. Однако в некоторых случаях оценка признаков по наилучшему типу распределения не дает интерпретируемого результата (все вероятности для 9 классов ниже порога значимости 0,15), как, например, получилось для раздела D, что может лишь свидетельствовать о высокой степени неопределенности или неустойчивости функционирования данного элемента подсистемы. Действительно, такой результат связан с тем, что в процедуре оценки вероятности на сопряженных числовой и лингвистической шкалах тип распределения остается неизменным, но меняются его числовые характеристики, например, математическое ожидание для нормального распределения — это среднее значение для интервала (класса в случае лингвистической шкалы). То есть для каждого класса меняются числовые характеристики для типа распределения, но не сам тип. При этом может возникать ситуация, когда значения вероятностей, рассчитанные для второго слагаемого в формуле (12), оказываются близкими друг другу, что и приводит к тому, что вероятности для всех классов не превышают порога значимости. Это накладывает ограничение на использование предложенного алгоритма, а выявленные случаи требуют более глубокого изучения и идентификации соответствующих причин. На данном этапе исследования можно рекомендовать использовать альтернативный тип распределения (со значимым и ближайшим к наилучшему типу распределения p-value) при получении вероятностных оценок.

На рис. 4 и рис. 5 представлена динамика объема ВРП для раздела С Тульской области за 2017–2023 гг., рассчитанная с использованием нормального и гамма-распределений. Модель нижнего уровня характеризует значение показателя, для которого, начиная с наименьшего значения, вероятность превышает порог 0,15. Наиболее вероятная модель характеризует значения показателя с наибольшей вероятностью. Модель верхнего уровня характеризует значение показателя, для которого, начиная с наибольшего значения, вероятность превышает порог 0,15.

Из рис. 5 видно, что модели нижнего, наиболее вероятного и верхнего уровней «схлопываются», уменьшая тем самым уровень неопределенности показателя.

Заключение

В результате исследования на основе байесовских интеллектуальных технологий была представлена методология оценки признаков, характеризующих результаты функционирования структурных элементов сложной системы с учетом наилучших типов распределений, соответствующих им при наличии тренда в данных. Представлены некоторые результаты апробации методологии на данных Тульской области.

Показано, что в некоторых случаях использование типа распределения при получении апостериорных оценок показателей в общем контуре байесовских интеллектуальных технологий приводит к уменьшению уровня их неопределенности, а также смещает итоговые оценки относительно нормативного значения по сравнению с оценками в предположении о нормальности типа распределения показателей. Это дает возможность улучшить обоснованность полученных результатов за счет использования дополнительной информации о характере и свойствах изучаемых процессов и структурных элементов систем.

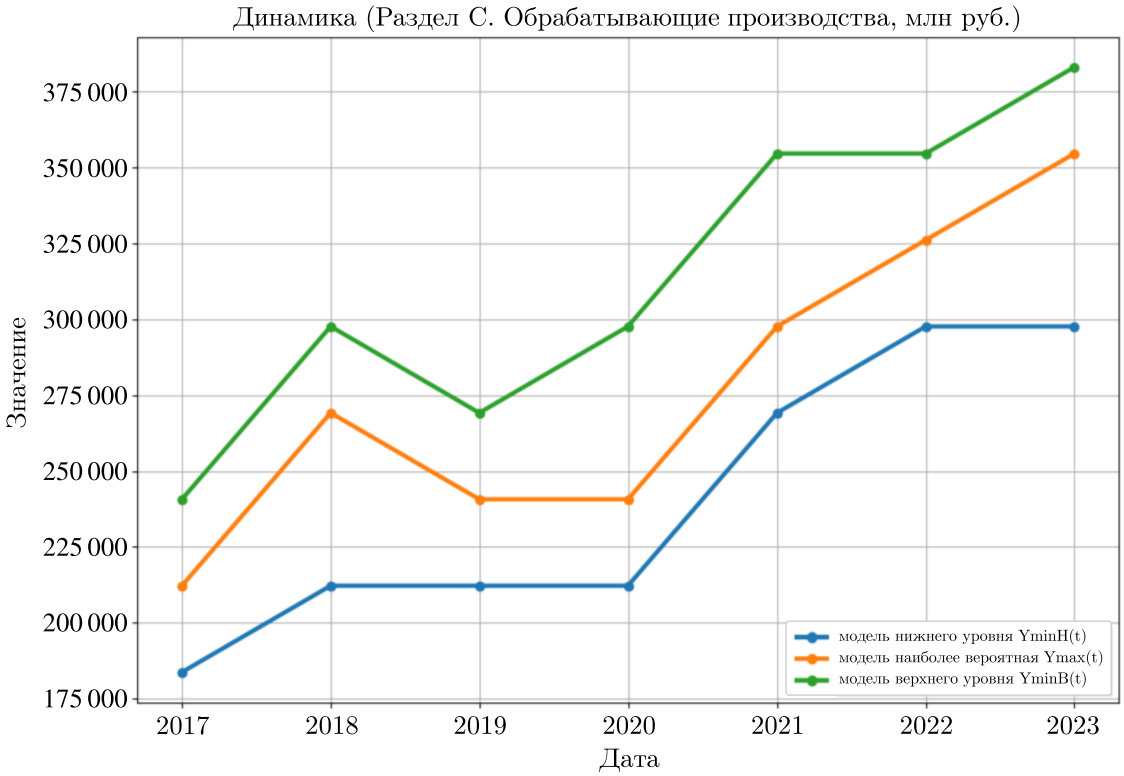

Рис. 4. Динамика объема ВРП для раздела C. Обрабатывающие производства Тульской области в 2017–2023 гг. при использовании альтернативного (нормального) распределения

Рис. 5. Динамика объема ВРП для раздела C. Обрабатывающие производства Тульской области в 2017–2023 гг. при использовании наилучшего (гамма) распределения

Результаты исследования могут быть использованы не только в региональных исследованиях, но и при изучении других сложных систем.

Список литературы (References)

Айвазян С. А., Мхитарян В. С. Прикладная статистика и основы эконометрики. — М.: ЮНИТИ, 1998. — 1022 с.

Ayvazyan S. A., Mkhitaryan V. S. Prikladnaya statistika i osnovy ekonometriki [Applied statistics and fundamentals of econometrics]. — Moscow: UNITY, 1998. — 1022 p. (in Russian).

Жуков Р. А. Моделирование системной сбалансированности региональной экономики в условиях неопределенности // МИР (Модернизация. Инновации. Развитие). — 2025a. — Т. 16, № 2. — С. 255–276.

Zhukov R. A. Modelirovaniye sistemnoy sbalansirovannosti regional'noy ekonomiki v usloviyakh neopredelennosti [Modeling the systemic balance of the regional economy under conditions of uncertainty] // MIR (Modernizatsiya. Innovatsii. Razvitiye) [MIR (Modernization. Innovation. Research)]. — 2025a. — Vol. 16, No. 2. — P. 255–276 (in Russian).

Жуков Р. А. О формировании норм на основе регрессионных моделей по малым выборкам // Чебышевский сборник. — 2025b. — Т. 26, № 1 (97). — С. 149–156.

Zhukov R. A. O formirovanii norm na osnove regressionnykh modeley po malym vyborkam [On the formation of norms based on regression models for small samples] // Chebyshevskii sbornik. — 2025b. — Vol. 26, No. 1 (97). — P. 149–156 (in Russian).

Жуков Р. А., Прокопчина С. В. О формировании мягких норм для оценки функционирования сложных систем // Чебышевский сборник. — 2024a. — Т. 25, № 3 (94). — С. 351–358.

Zhukov R. A., Prokopchina S. V. O formirovanii myagkikh norm dlya otsenki funktsionirovaniya slozhnykh sistem [On the formation of soft norms for evaluating the functioning of complex systems] // Chebyshevskii sbornik. — 2024a. — Vol. 25, No. 3 (94). — P. 351–358 (in Russian).

Жуков Р. А., Прокопчина С. В. Свидетельство о государственной регистрации программы для ЭВМ № 2024617544 Российская Федерация. Программный комплекс «Инфоаналитик 2.0»: № 2024615345: заявл. 10.03.2024: опубл. 03.04.2024b.

Zhukov R. A., Prokopchina S. V. Svidetel'stvo o gosudarstvennoy registratsii programmy dlya EVM No. 2024617544 Rossiyskaya Federatsiya. Programmnyy kompleks "Infoanalitik 2.0": No. 2024615345: zayavl. 10.03.2024: opubl. 03.04.2024b [Certificate of state registration of computer program No. 2024617544 Russian Federation. Software package "Infoanalitik 2.0": No. 2024615345: declared 10.03.2024: published 03.04.2024b] (in Russian).

Жуков Р. А., Прокопчина С. В., Гиниатов И. А., Николина Е. М. Применение библиотеки «Байесовская математическая статистика» в программном комплексе «Инфоинтегратор» // Мягкие измерения и вычисления. — 2022. — Т. 54, № 5. — С. 99–108.

Zhukov R. A., Prokopchina S. V., Giniatov I. A., Nikolina E. M. Primeneniye biblioteki "Bayyesovskaya matematicheskaya statistika" v programmnom komplekse "Infointegrator" [Application of the Bayesian Mathematical Statistics library in the Infointegrator software package] // Myagkiye izmereniya i vychisleniya [Soft Measurements and Computing]. — 2022. — Vol. 54, No. 5. — P. 99–108 (in Russian).

Журавлев С. Д., Жуков Р. А. Методика анализа эффективности государственного управления (на примере регионов РФ) // Известия Пензенского государственного педагогического университета им. В. Г. Белинского. — 2012. — № 28. — С. 349–357.

Zhuravlev S. D., Zhukov R. A. Metodika analiza effektivnosti gosudarstvennogo upravleniya (na primere regionov RF) [The methods of state management effectiveness analysis (on the example of the RF regions)] // Izvestiya Penzenskogo gosudarstvennogo pedagogicheskogo universiteta im. V. G. Belinskogo [Bulletin of the Penza State Pedagogical University named after V. G. Belinsky]. — 2012. — No. 28. — P. 349–357 (in Russian).

Инфляция в России. — [Электронный ресурс]. — https:// уровень-инфляции.рф/таблицы-инфляции (дата обращения: 11.01.2026).

Inflyatsiya v Rossii [Inflation in Russia]. — [Electronic resource]. — https://xn—-ctbjnaatncev9av3a8f8b.xn–p1ai/ (accessed: 11.01.2026).

Клейнер Г. Б. Флагман экономико-математического и компьютерного моделирования: 60 лет в строю // Экономика и математические методы. — 2023. — Т. 59, № 3. — С. 5–20.

Kleiner G. B. Flagman ekonomiko-matematicheskogo i komp'yuternogo modelirovaniya: 60 let v stroyu [The flagship of economic, mathematical and computer modeling: 60 years in line] // Ekonomika i matematicheskiye metody [Economics and Mathematical Methods]. — 2023. — Vol. 59, No. 3. — P. 5–20 (in Russian).

Клейнер Г. Б., Рыбачук М. А. Системная сбалансированность экономики: монография. — М.: Научная библиотека, 2017. — 320 с.

Kleiner G. B., Rybachuk M. A. Sistemnaya sbalansirovannost' ekonomiki: monografiya [System balance of the economy: Monograph]. — Moscow: Nauchnaya biblioteka [Scientific Library Publ], 2017. — 320 p. (in Russian).

Клейнер Г. Б., Рыбачук М. А. Системная сбалансированность экономики России: региональный разрез // Экономика региона. — 2019. — Т. 15, № 2. — С. 309–323.
Kleiner G. B., Rybachuk M. A. Sistemnaya sbalansirovannost' ekonomiki Rossii: regional'nyy razrez [System balance of the Russian economy: regional perspective] // Ekonomika regiona [Economy of Regions]. — 2019. — Vol. 15, No. 2. — P. 309–323 (in Russian).

Лемешко Б. Ю. Непараметрические критерии согласия. Руководство по применению: монография. — М.: ИНФРА-М, 2024. — 201 с.
Lemeshko B. Yu. Neparametricheskiye kriterii soglasiya. Rukovodstvo po primeneniyu: monografiya [Nonparametric Criteria of Agreement. Application Guide: Monograph]. — Moscow: INFRA-M, 2024. — 201 p. (in Russian).

Прокопчина С. В. Байесовские интеллектуальные измерения. — М.: Научная библиотека, 2021a. — 495 с.
Prokopchina S. V. Bayyesovskiye intellektual'nyye izmereniya [Bayesian intellectual measurements]. — Moscow: Nauchnaya biblioteka [Scientific Library Publ.], 2021a. — 495 p. (in Russian).

Прокопчина С. В. Байесовские интеллектуальные технологии в задачах моделирования закона распределения в условиях неопределенности. — М.: Научная библиотека, 2020. — 292 с.
Prokopchina S. V. Bayyesovskiye intellektual'nyye tekhnologii v zadachakh modelirovaniya zakona raspredeleniya v usloviyakh neopredelennosti [Bayesian intelligent technologies in problems of modeling the distribution law under uncertainty]. — Moscow: Nauchnaya biblioteka [Scientific Library Publ.], 2020. — 292 p. (in Russian).

Прокопчина С. В. Основы теории шкалирования в экономике: учебное пособие. — М.: Научная библиотека, 2021b. — 272 с.
Prokopchina S. V. Osnovy teorii shkalirovaniya v ekonomike: uchebnoye posobiye [Fundamentals of scaling theory in economics: a textbook]. — Moscow: Nauchnaya biblioteka [Scientific Library Publ.], 2021b. — 272 p. (in Russian).

Регионы России. Социально-экономические показатели. Статистический сборник. 2020–2024 гг. — [Электронный ресурс]. — https://rosstat.gov.ru/folder/210/document/13204 (дата обращения: 11.01.2026).
Regiony Rossii. Sotsial'no-ekonomicheskiye pokazateli. Statisticheskiy sbornik. 2020–2024 gg. [Regions of Russia. Socio-economic indicators. Statistical collection. 2020–2024]. — [Electronic resource]. — https://rosstat.gov.ru/folder/210/document/13204 (accessed: 11.01.2026).

PyCharm IDE для профессиональной разработки на Python. — [Электронный ресурс]. — https://www.jetbrains.com/ru-ru/pycharm (дата обращения: 11.01.2026).

Akaike H. A new look at the statistical model identification // IEEE Transactions on Automatic Control. — 1974. — Vol. 19. — P. 716–723.

Angelini O., Gabrielli A., Tacchella A., Zaccaria A., Pietronero L., Matteo T. Di. Forecasting the countries' gross domestic product growth: The case of Technological Fitness // Chaos, Solitons & Fractals. — 2024. — Vol. 184. — 115006.

Charpentier A., Flachaire E. Pareto models for top incomes and wealth // The Journal of Economic Inequality. — 2022. — Vol. 20, No. 1. — P. 1–25.

Iacopini M., Poon A., Rossini L., Zhu D. Bayesian mixed-frequency quantile vector autoregression: Eliciting tail risks of monthly US GDP // Journal of Economic Dynamics & Control. — 2023. — Vol. 157. — P. 104757.

Koop G., Korobilis D. Bayesian dynamic variable selection in high dimensions // International Economic Review. — 2023. — Vol. 64, No. 3. — P. 1047–1074.

Mirzaei Sh., Jahanshahi S. M. A. A new goodness of fit measure based on income inequality curves // Journal of Modern Applied Statistical Methods. — 2022. — Vol. 19, No. 1. — Article 24.

Ramos A., Massing T., Ishikawa A., Fujimoto S., Mizuno T. Mixtures of log-normal distributions in the mid-scale range of firm-size variables // Evolut. Inst. Econ. Rev. — 2024. — Vol. 21. — P. 249–260.

Silbersdorff A., Schneider K. S. Distributional regression techniques in socioeconomic research on the inequality of health with an application on the relationship between mental health and income // Int. J. Environ. Res. Public Health. — 2019. — Vol. 16, No. 20. — P. 4009.

Taleb N. N. Statistical consequences of fat tails: real world preasymptotics, epistemology, and applications. — Third edition. — STEM Academic Press, 2025. — 523 p.