

# Некоторые вопросы измерений и вычислений при управлении сложными системами

Р. А. Жуков<sup>1</sup>, С. В. Прокопчина<sup>2</sup>, М. А. Плинская<sup>1</sup>, М. А. Желуницина<sup>1</sup>

<sup>1</sup>Финансовый университет при Правительстве Российской Федерации (Тульский филиал)

<sup>2</sup>Финансовый университет при Правительстве Российской Федерации

pluszh@mail.ru, svprokopchina@mail.ru, maria.plinskaya@gmail.com, maria202001@yandex.ru

**Аннотация.** Решается задача формирования нормативов и процедур вычисления индикаторов функционирования сложных систем в условиях неопределенности. В основе методологии лежит методология байесовских интеллектуальных технологий, дающая возможность осуществлять моделирование в условиях ограниченности, неполноты и нечеткости данных, в том числе при единичных измерениях. Разработанная методика включает: представление измерений в виде нечеткого числа; идентификацию типа распределения; генерацию выборок заданного объема для факторов, характеризующих результат и условия функционирования сложной системы в заданный период времени; построение регрессионной модели по сформированным выборкам. Построенные модели могут выступать в качестве инструмента для вычисления динамических нормативов посредством подстановки в уравнение соответствующих факторов.

**Ключевые слова:** система; индикатор; методология байесовских интеллектуальных измерений

## I. ВВЕДЕНИЕ

Применение методологии байесовских интеллектуальных технологий (БИТ) [1–3] при решении задач управления сложными системами, в частности социо-эколого-экономическими системами (СЭЭС), дает возможность работать в условиях неопределенности, неполноты и нечеткости данных, характерных для систем подобного типа. В силу ограниченности информации о поведении объекта в определенный период времени возникает проблема формирования достаточной выборки для обеспечения корректной, с точки зрения математической статистики, процедуры построения адекватных регрессионных моделей. В эконометрических исследованиях вводят допущения – предположения о том, что объем выборки достаточен для конструирования таких моделей, используемых для оценки и анализа сложных систем (СЭЭС) [4–6], в том числе посредством частных и интегральных индикаторов [8–10]. При этом можно говорить, что измерение в определенный период времени для подсистемы или элемента СЭЭС является единичным, уникальным в части условий проведения наблюдения, что также влияет на объективность получаемых результатов. В этом аспекте актуальным является формирование алгоритма построения регрессионных моделей для элемента, подсистемы или системы в целом, характеризующихся

одним или несколькими частными (для элемента), интегральными (для подсистемы, системы) результативными признаками и условиями – факторами – их функционирования в конкретный период времени. Результаты расчета по модели могут выступать как вариант норматива для выделенного результативного признака, характеризующего результат функционирования СЭЭС, ее подсистем и элементов.

Таким образом, целью исследования является разработка методики построения регрессионных моделей, используемых для формирования нормативов по малым выборкам, в том числе на базе единичных измерений.

## II. МЕТОДОЛОГИЯ ИССЛЕДОВАНИЯ

Методология исследования базируется на БИТ [1–3, 11] в рамках регуляризирующего байесовского подхода, который предполагает представление результата измерения в виде нечеткого числа на заданной шкале.

Результат измерения представляется в виде набора гипотез, характеризующихся значением случайной величины и вероятностей появления гипотез, вычисленных по формуле Байеса:

$$P(H_i | S) = P(H_i)P(S | H_i) / \sum_{i=1}^k P(H_i)P(S | H_i), \quad (1)$$

где  $H_i$  –  $i$ -е эталонное распределение;  $P(H_i)$  – априорная вероятность появления  $H_i$ -го эталонного распределения;  $S$  – событие, заключающееся в совместном появлении значений оценок индикатора;  $P(S|H_i)$  – апостериорная вероятность совместного появления значений индикаторов при условии справедливости гипотезы  $H_i$ ;  $k$  – количество классов эталонных гипотез.

Преобразованные таким образом значения результата измерений для заданного объекта (СЭЭС, ее подсистем элементов) является входом – первым этапом алгоритма.

На втором этапе осуществляется идентификация типа распределения посредством определения разделяющих границ на плоскости моментов Пирсона [12] по алгоритму, реализованному в программном модуле [13].

Третий этап предполагает генерацию требуемого объема выборки с идентифицированным на предыдущем шаге типом распределения для каждого из набора результативных и факторных признаков.

Четвертый этап включает построение регрессионных моделей связи результативных и факторных признаков

Исследование выполнено за счет гранта Российского научного фонда 24-28-20020, <https://rscf.ru/project/24-28-20020/> и Тульской области

по достаточному объему выборки с точки зрения математической статистики.

### III. ИНФОРМАЦИОННАЯ БАЗА И РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Для реализации предложенной методики были использованы открытые данные Федеральной службы государственной статистики Российской Федерации для Тульской области [14] в 2022 г.

В качестве результативного признака использован объем валового регионального продукта (ВРП) по виду деятельности «Обрабатывающие производства» (раздел D), а в качестве факторных признаков выбраны: стоимость основных производственных фондов и среднегодовая численность занятых по соответствующему виду деятельности. Стоимостные факторы скорректированы на уровень инфляции и приведены к уровню 2007 года. Соответствующие фактические значения в 2022 году составили (для раздела D): ВРП – 107512.580 млн руб.; стоимость основных фондов – 181342.720 тыс. руб.; среднегодовая численность занятых – 161.100 тыс. чел.

Для шкалирования была использована базовая реализация программной платформы «Инфоаналитик» [15] на языке Python «Инфоаналитик 2.0». На первом этапе в качестве норм были выбраны средние значения для результативных и факторных признаков по данным за 2007–2022 гг.

Лингвистическая шкала для факторов представлена на рис. 1.

Для каждого из факторов были сгенерированы наборы данных объемом 100, 1000 и 7000 в соответствии с алгоритмом [13]. Для каждого репера  $x_i$  числовой шкалы моделировалось треугольное распределение на интервале  $[x_i - (x_i - x_{i-1})/2; x_i + (x_{i+1} - x_i)/2]$ . Число реперов 9. Тестирование показало, что полученные выборки соответствуют нормальному распределению, как показано на рис. 2 для выборки объемом 7000. Оценка показала, что случайные величины соответствуют нормальному закону распределения.

Полученные данные были использованы для построения степенных мультипликативных моделей:

$$y_{100} = 28.883 \cdot x_1^{0.100} \cdot x_2^{1.381}, R^2 = 0.931, \quad (2)$$

$$y_{1000} = 15.446 \cdot x_1^{0.172} \cdot x_2^{1.330}, R^2 = 0.920, \quad (3)$$

$$y_{7000} = 16.665 \cdot x_1^{0.164} \cdot x_2^{1.335}, R^2 = 0.917. \quad (4)$$

Из формул (2)–(4) видно, что коэффициент детерминации для моделей довольно высокий.

Средние относительные ошибки (MAPE) составили 3.021 %, 3,250 % и 3.220 % соответственно.

Расчетные значения валового регионального продукта по моделям, построенным по выборкам объемов 100, 1000 и 7000 соответственно, составили: 107821.911, 107050.080 и 107179.403 тыс. руб. В первом приближении расчетные значения могут выступать в качестве норматива для оценки результата функционирования обрабатывающей отрасли в Тульской области.

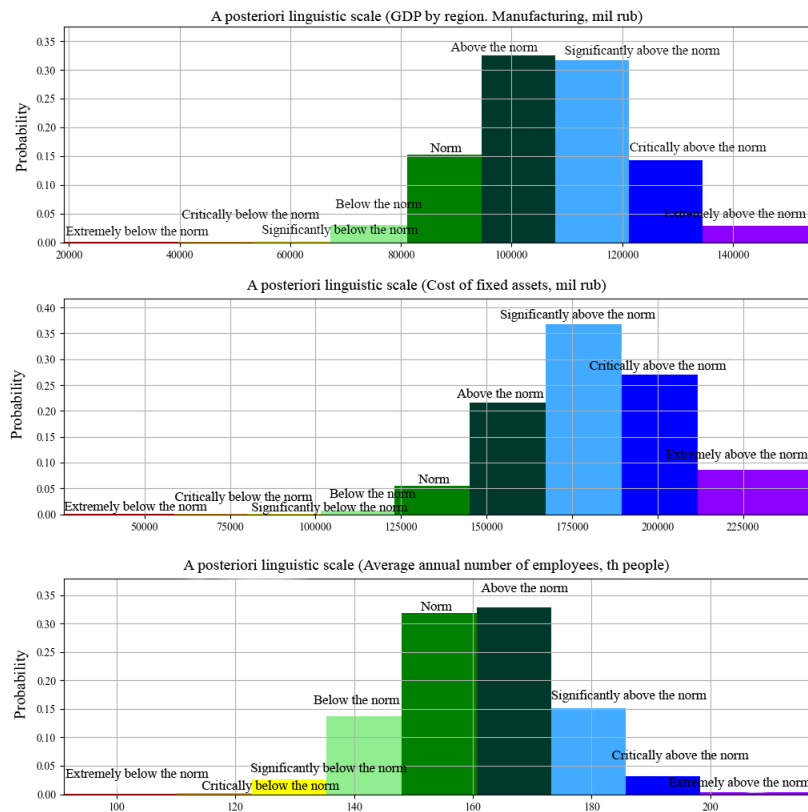
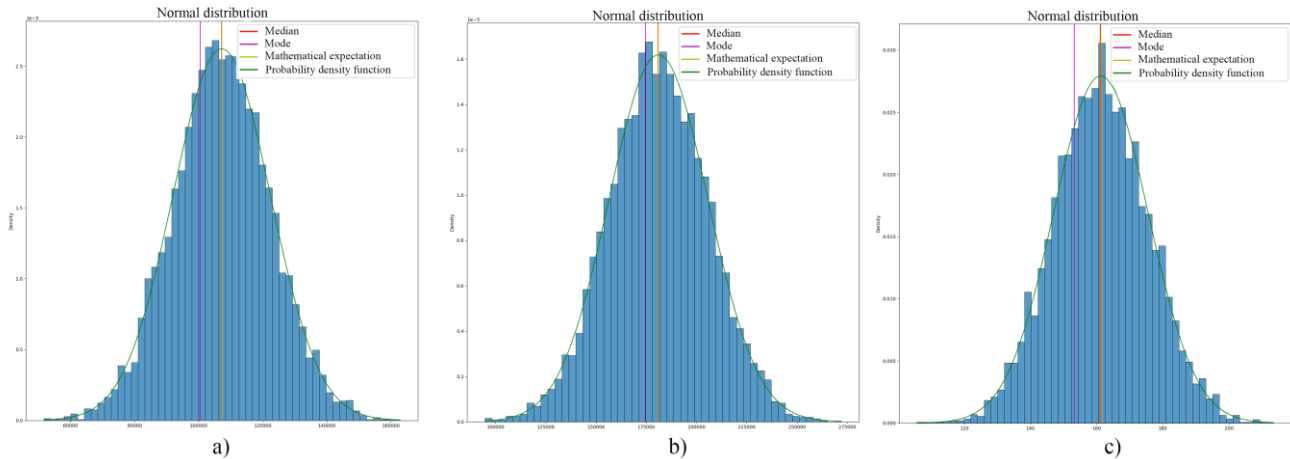


Рис. 1. Лингвистическая апостериорная шкала факторов



Примечание: а) объем ВРП; б) стоимость основных фондов; с) среднегодовая численность занятых.

Рис. 2. Результат генерации данных объемом 7000 по малым выборкам

#### IV. ЗАКЛЮЧЕНИЕ

Представлена методика построения регрессионных моделей, используемых для формирования нормативов по малым выборкам, в том числе на базе единичных измерений. Материалы исследования могут быть использованы для построения моделей функционирования сложных систем (СЭЭС) при ограниченности исходных данных. Модели могут служить основой формирования управленческих решений посредством вариации факторных признаков и разработки соответствующих мероприятий.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Прокопчина С.В., Щербаков Г.А., Ефимов Ю.В. Моделирование социально-экономических систем в условиях неопределенности: учебное пособие. М.: Научная библиотека, 2019. 508 с.
- [2] Прокопчина С.В. Мягкие вычисления и измерения. Методы, информационные технологии и средства интеллектуальной обработки информации в задачах цифровизации: монография / под ред. д.т.н., проф. С.В. Прокопчиной. М.: ИД «НАУЧНАЯ БИБЛИОТЕКА», 2019. 616 с.
- [3] Прокопчина С.В. Основы теории шкалирования в экономике: учебное пособие. М.: Научная библиотека, 2021. 298 с.
- [4] Zhukov R.A., Kozlova N.A., Manokhin E.V., Plinskaya M.A. Construction of an aggregated production function with implementation based on the example of the regions of the Central federal district of the Russian Federation // *Business Informatics*. 2022. Vol. 16. No. 3. Pp. 7-23.
- [5] Wang L., Rang X., Mu L. The Coupling Coordination Evaluation of Sustainable Development between Urbanization, Housing Prices, and Affordable Housing in China // *Discrete Dynamics in Nature and Society*. Vol. 2021. Article ID 3937226.
- [6] Pedroza C. A Bayesian forecasting model: predicting U.S. male mortality // *Biostatistics*. 2006. Vol. 7. No. 4. Pp. 530–550.
- [7] Жуков Р.А. Некоторые задачи оптимизации управления социально-экономическими системами // *Чебышевский сборник*. 2019. Т. 20. № 1 (69). С. 370-388.
- [8] Bormotov I.V., Prokopchina S.V., Zhukov R.A. Methodology of Bayesian intellectual measurements for information systems and its application to the study of the value system of russian youth // *XXVI International Conference on Soft Computing and Measurements. Proceedings of 2023 XXVI International Conference on Soft Computing and Measurements (SCM)*. 2023. Pp. 7-10.
- [9] Березин А.С., Жуков Р.А., Прокопчина С.В. Байесовские интеллектуальные измерения индексов и показателей региональной обеспеченности объектами культуры // *Мягкие измерения и вычисления*. 2022. Т. 53. № 4. С. 5-15.
- [10] Zhukov R.A. An approach to assessing the functioning of hierarchical socio-economic systems and decision-making based on the EFRA software package // *Business Informatics*. 2020. Vol. 14. No. 3. Pp.82-95.
- [11] Прокопчина С.В. Байесовские интеллектуальные измерения. М.: Издательский дом «НАУЧНАЯ БИБЛИОТЕКА», 2021. 495 с.
- [12] Прокопчина С.В. Байесовские интеллектуальные технологии в задачах моделирования закона распределения в условиях неопределенности: монография. М.: Издательский дом «НАУЧНАЯ БИБЛИОТЕКА», 2020. 292 с.
- [13] 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.
- [14] Федеральная служба государственной статистики Российской Федерации. Официальный сайт. URL: <https://rosstat.gov.ru/folder/210/document/13204>.
- [15] Прокопчина С.В. Инфоаналитик (свидетельство Федеральной Службы по интеллектуальной собственности, патентам и торговым знакам об официальной регистрации программы для ЭВМ № 2004611741 от 12.08.2004).