

Трусова, А. Ю. Динамический и канонический анализы как инструмент прогнозирования посещаемости амбулаторно-поликлинических организаций в Самарской области / А. Ю. Трусова, А. И. Ильина // Прикладная математика и вопросы управления. – 2025. – № 1. – С. 59–83. DOI 10.15593/2499-9873/2025.1.05

Библиографическое описание согласно ГОСТ Р 7.0.100–2018

Трусова, А. Ю. Динамический и канонический анализы как инструмент прогнозирования посещаемости амбулаторно-поликлинических организаций в Самарской области / А. Ю. Трусова, А. И. Ильина. – Текст : непосредственный. – DOI 10.15593/2499-9873/2025.1.05 // Прикладная математика и вопросы управления / Applied Mathematics and Control Sciences. – 2025. – № 1. – С. 59–83.



ПРИКЛАДНАЯ МАТЕМАТИКА
И ВОПРОСЫ УПРАВЛЕНИЯ

№ 1, 2025

<https://ered.pstu.ru/index.php/amcs>



Научная статья

DOI: 10.15593/2499-9873/2025.1.05

УДК 330.43



Динамический и канонический анализы как инструмент прогнозирования посещаемости амбулаторно-поликлинических организаций в Самарской области

А.Ю. Трусова, А.И. Ильина

Самарский университет, Самара, Российская Федерация

О СТАТЬЕ

Получена: 12 ноября 2024
Одобрена: 03 марта 2025
Принята к публикации:
18 апреля 2025

Финансирование

Исследование не имело спонсорской поддержки.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Вклад авторов

равноценен.

Ключевые слова:

динамический анализ, канонический анализ, оценка прогнозного уровня, сглаживание простой скользящей средней, гипотезы о наличии тренда, трендовые модели, модели авторегрессии и скользящего среднего, канонический коэффициент корреляции, канонические переменные

АННОТАЦИЯ

Исследование описывает результаты применения динамического и канонического анализов при прогнозировании посещаемости амбулаторно-поликлинических организаций в Самарской области. Сфера здравоохранения в настоящий период находится в сложном состоянии. Это можно рассматривать как с точки зрения собственно медицинских учреждений, так и их посетителей. В системе здравоохранения амбулаторно-поликлинические учреждения играют ключевую роль. Следовательно, разработка и внедрение инновационных подходов к управлению, улучшению качества медицинского обслуживания и повышению доступности медицинских услуг в амбулаторно-поликлинических учреждениях является одной из главных задач современной системы здравоохранения. Для анализа посещаемости медицинских учреждений существует множество статистических методов, включая динамический и канонический анализ. Именно в комплексном подходе к анализу с использованием перечисленных методов и заключается научная новизна работы. Это позволяет выявить не только прямые, но и косвенные зависимости между численностью посетителей и состоянием окружающей среды. Практическая значимость исследования выражается в возможности использования полученных результатов для корректировки стратегий управления и развития амбулаторно-поликлинических учреждений, улучшения качества медицинского обслуживания и повышения эффективности использования ресурсов здравоохранения. Целью данной работы является оценка прогнозного числа посещаемости амбулаторно-поликлинических организаций в Самарской области за период с 2004 по 2021 г., а также исследование степени взаимосвязи с экологическими факторами.

© Трусова Алла Юрьевна – кандидат физико-математических наук, доцент, доцент кафедры математики и бизнес-информатики, e-mail: a_yu_ssu@mail.ru, ORCID 0000-0001-7679-9902.

Ильина Алла Ивановна – старший преподаватель кафедры математики и бизнес-информатики, e-mail: ilina.ai@ssau.ru, ORCID <https://orcid.org/0000-0002-7624-5771>.



Эта статья доступна в соответствии с условиями лицензии Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

Perm Polytech Style: Trusova A.Yu., Ilyina A.I. Dynamic and canonical analysis as a tool for predicting attendance of outpatient clinics in the Samara region. *Applied Mathematics and Control Sciences*. 2025, no. 1, pp. 59–83. DOI: 10.15593/2499-9873/2025.1.05

MDPI and ACS Style: Trusova, A.Yu.; Ilyina, A.I. Dynamic and canonical analysis as a tool for predicting attendance of outpatient clinics in the Samara region. *Appl. Math. Control Sci.* **2025**, *1*, 59–83. <https://doi.org/10.15593/2499-9873/2025.1.05>

Chicago/Turabian Style: Trusova, Alla Yu., and Alla I. Ilyina. 2025. “Dynamic and canonical analysis as a tool for predicting attendance of outpatient clinics in the Samara region”. *Appl. Math. Control Sci.* no. 1: 59–83. <https://doi.org/10.15593/2499-9873/2025.1.05>



APPLIED MATHEMATICS
AND CONTROL SCIENCES

№ 1, 2025

<https://ered.pstu.ru/index.php/amcs>



Article

DOI: 10.15593/2499-9873/2025.1.05

UDC 330.43



Dynamic and canonical analysis as a tool for predicting attendance of outpatient clinics in the Samara region

A.Yu. Trusova, A.I. Ilyina

Samara University, Samara, Russian Federation

ARTICLE INFO

Received: 12 November 2024

Approved: 03 March 2025

Accepted for publication:

18 April 2025

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

equivalent.

Keywords:

dynamic analysis, canonical analysis, forecast level assessment, simple moving average smoothing, trend hypotheses, trend models, autoregressive and moving average models, canonical correlation coefficient, canonical variables

ABSTRACT

The study describes the results of using dynamic and canonical analysis in forecasting the attendance of outpatient clinics in the Samara region. The healthcare sector is currently in a difficult state. This can be considered both from the point of view of the medical institutions themselves and their visitors. Outpatient clinics play a key role in the healthcare system. Therefore, the development and implementation of innovative approaches to management, improving the quality of medical care and increasing the availability of medical services in outpatient clinics is one of the main tasks of the modern healthcare system. There are many statistical methods for analyzing the attendance of medical institutions, including dynamic and canonical analysis. It is in the integrated approach to analysis, using the listed methods, that the scientific novelty of the work lies. This allows us to identify not only direct but also indirect relationships between the number of visitors and the state of the environment. The practical significance of the study is expressed in the possibility of using the obtained results to adjust management strategies and development of outpatient clinics, improve the quality of medical care and increase the efficiency of using healthcare resources. The purpose of this work is to assess the predicted number of visits to outpatient clinics in the Samara region for the period from 2004 to 2021, as well as to study the degree of relationship with environmental factors.

© Alla Yu. Trusova – CSc in Physics and Mathematics, Associate Professor, Associate Professor of the Department of Mathematics and Business Informatics, e-mail: a_yu_ssu@mail.ru, ORCID 0000-0001-7679-9902.

Alla Ia. Ilyina – Senior Lecturer of the Department of Mathematics and Business Informatics, e-mail: ilina.ai@ssau.ru, ORCID: 0000-0002-7624-5771.



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

Введение

Современное состояние здравоохранения характеризуется в настоящий период как сложное с позиции и медицинских учреждений, и их посетителей. Проблемы медицинского обслуживания и развития сферы медицинских услуг достаточно остро обсуждаются в научной литературе. Исследователи с разных сторон описывают проблемы и предлагают варианты решения. Так, автор [1] глубоко анализирует качество жизни населения Красноярского края, выявляет факторы, влияющие на улучшение жизни граждан. В исследовании используется широкий диапазон показателей социально-экономической сферы. Автор [2] в своем исследовании представляет метод информационного обеспечения принятия управленческих решений в сфере здравоохранения на основе математического экспресс-анализа статистических показателей заболеваемости населения, которые рассмотрены в сравнении по территориям и периодам времени. На уровне федерального округа в сфере здравоохранения выявлены проблемы, решение которых значительно повысит качество медицинских услуг, оказываемых населению. Результаты исследования [3] могут быть использованы в области общественного здравоохранения. Особо отмечается в работе связь со здравоохранением, а именно: как последствия перехода к «обществу долголетия» влияют на структуру заболеваемости, ожидаемую продолжительность здоровой жизни, финансовые расходы на отдельные виды медицинской помощи, зависимость населения от медицинской помощи. Вопросам лекарственного обеспечения посвящено исследование [4]. В результате экспертного опроса специалистов разработаны приоритетные направления совершенствования лекарственного обеспечения в Арктике, которые могут быть рекомендованы для использования при формировании концепции развития фармацевтической службы арктических территорий, при создании общей картины идей, связанных с поиском путей решения имеющихся проблем. Работа [5] представляет результаты исследования по совершенствованию методологии и практических механизмов реализации оценки эффективности управления бюджетными ресурсами в сфере здравоохранения. С помощью методов математической статистики, экономического и системного анализа решаются вопросы оценки эффективности управления бюджетными ресурсами в сфере здравоохранения. Применение аппарата имитационного моделирования для оптимизации работы медицинских организаций разных уровней и направленности представлено в работе [6]. Автор обращает внимание, что именно имитационное моделирование позволяет оценить нагрузку как на отдельные этапы, так и на все эвакуационное направление до непосредственного задействования сил и средств медицинской службы. В работе предложена система анализа и визуализации данных, используемых для оптимизации принятия решения о распределении сил и средств медицинской службы на эвакуационном направлении с применением имитационного моделирования. Автором [7] представлены результаты применения статистического анализа временных рядов с использованием линии тренда, результаты применения регрессионного анализа и прогнозирования при изучении показателей медицинской сферы. В работе [8] представлен обширный математический инструментарий при анализе качества жизни населения. Кроме описанного, в работах [9–15] приводятся результаты применения многомерных методов при изучении вопросов качества жизни и их взаимосвязь с медицинской сферой. Данные официальной статистики рассматриваются в работах [16–19]. Обобщая, следует отметить, что актуальность вопросов развития здравоохранения в России прослеживается авторами всех работ. Дополнительно следует отметить, что в системе здравоохранения амбулаторно-поликлинические учреждения играют ключевую

роль. Они служат важным звеном в цепи медицинской помощи. Следовательно, разработка и внедрение инновационных подходов к управлению, улучшению качества медицинского обслуживания и повышению доступности медицинских услуг в амбулаторно-поликлинических учреждениях в настоящем периоде является одной из главных задач современной системы здравоохранения.

Роль и значение амбулаторно-поликлинических учреждений в системе здравоохранения определяют актуальность данного исследования. Растущий спрос на медицинские услуги, изменение в структуре заболеваемости, а также необходимость адаптации системы здравоохранения к динамично меняющимся условиям внешней среды определяют актуальность данной работы.

Для анализа посещаемости медицинских учреждений существует множество статистических методов, включая динамический и канонический анализ. Именно в комплексном подходе к анализу с использованием перечисленных методов и заключается научная новизна работы. Это позволяет выявить не только прямые, но и косвенные зависимости между посещаемостью и состоянием окружающей среды.

Практическая значимость исследования выражается в возможности использования полученных результатов для корректировки стратегий управления и развития амбулаторно-поликлинических учреждений, улучшения качества медицинского обслуживания и повышения эффективности использования ресурсов здравоохранения.

Целью данной работы является оценка прогнозного числа посещаемости амбулаторно-поликлинических организаций в Самарской области за период с 2004 по 2021 г., а также исследование степени взаимосвязи с экологическими факторами.

Поскольку спрос на медицинские услуги продолжает расти, центры амбулаторной помощи превратились в важнейшие компоненты системы здравоохранения, предлагая широкий спектр диагностических, терапевтических и профилактических услуг за пределами традиционных больничных учреждений.

Таким образом, можно выделить следующие преимущества амбулаторных клиник:

1) своевременность обслуживания, то есть пациенты могут получить необходимую медицинскую помощь в течение дня без предварительной записи на длительный срок, что особенно важно для людей с острыми состояниями или для тех, кто нуждается в регулярном медицинском наблюдении;

2) удобство, так как амбулаторные центры часто расположены ближе к месту жительства пациентов, что сокращает время и расходы на транспорт, а также упрощает процесс получения медицинской помощи;

3) комплексный подход – многие центры амбулаторной помощи предлагают широкий спектр услуг, включая диагностику, лечение, консультации специалистов и реабилитационные процедуры, что позволяет оказывать комплексную медицинскую помощь;

4) индивидуализированный уход – в центрах амбулаторной помощи часто уделяется больше внимания персональному подходу к каждому пациенту, учитывая его личные потребности и предпочтения;

5) профилактика и обучение, амбулаторные центры также играют важную роль в профилактике заболеваний и обучении пациентов основам здорового образа жизни, что способствует улучшению общего состояния здоровья населения.

Согласно статистике Росстата, по результатам 2021 г. в Российской Федерации наблюдается снижение числа амбулаторно-поликлинических учреждений на 1,3 тысячи, при этом в 2020 г. было зафиксировано 22,9 тысячи подобных организаций. В то же время, не-

смотря на уменьшение количества амбулаторий и поликлиник, их средняя посещаемость по стране увеличилась на 106 тысяч человек за смену в 2021 г. В частности, в Центральном федеральном округе пропускная способность этих учреждений возросла на почти 8 %, а в Дальневосточном федеральном округе – на 5 %. Сравнительная статистика представлена на рис. 1.

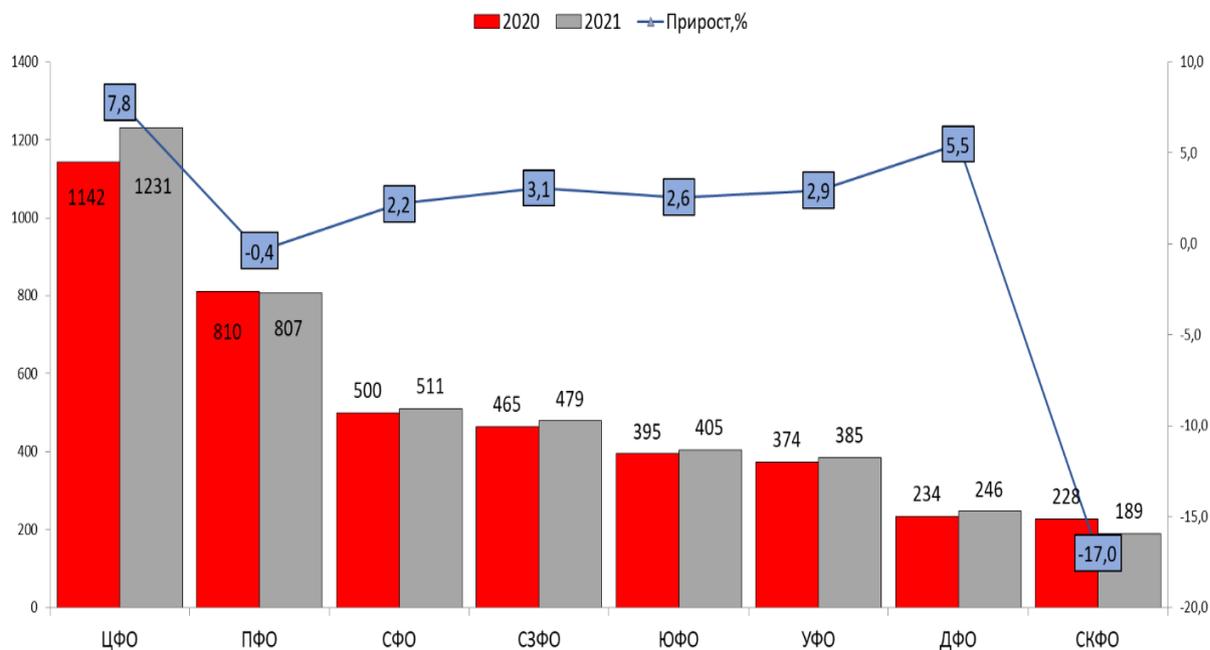


Рис. 1. Динамика мощности амбулаторно-поликлинических организаций в федеральных округах в 2020–2021 гг., тыс. посещений в смену

По данным медицинской статистики более половины платных медицинских услуг, предоставляемых гражданам России в медицинских учреждениях, приходится на амбулаторный сектор, составляя 58,9 % от общего объема медицинской помощи, финансируемой за счет средств населения.

Рост популярности амбулаторно-поликлинических организаций в последние годы стал заметным трендом, который отражает изменения в предпочтениях населения и стремление системы здравоохранения к оптимизации своих услуг.

Амбулаторно-поликлинические организации в Самарской области

В рамках российской медицинской инфраструктуры амбулаторно-поликлинические организации занимают центральное место, выполняя ключевую роль в обеспечении населения первичной и специализированной медицинской помощью. Эти учреждения служат первым и зачастую самым доступным контактом пациентов с системой здравоохранения, предлагая широкий спектр услуг – от профилактических мероприятий до диагностики и лечения различных заболеваний. Самарская область активно развивает свою сеть амбулаторно-поликлинических учреждений, стремясь обеспечить качественное и всестороннее медицинское обслуживание населения. Учреждения функционируют как в городских, так и в сельских районах, делая медицинские услуги более доступными для всех категорий граждан. Особое внимание в регионе уделяется развитию первичного звена здравоохранения, включая усиление роли семейных врачей и педиатров, которые играют ключевую роль в профилактике, раннем выявлении и управлении хроническими заболеваниями.

Важным аспектом работы амбулаторно-поликлинических организаций в Самарской области является стремление к улучшению доступности медицинских услуг для маломобильных групп населения и жителей отдаленных районов. Реализация программ дистанционного медицинского консультирования и телемедицины позволяет расширить возможности для своевременного и качественного медицинского обслуживания, минимизируя необходимость в поездках за медицинской помощью.

Согласно официальной статистике [20], в период с 2015 по 2019 г. наблюдалась неоднородная динамика в натуральном объеме медицинского рынка Самары. В 2016 г. произошло наиболее значительное уменьшение количества медицинских приемов на 3,4 %, по сравнению с предыдущим годом, вызванное ухудшением доступности медицинских услуг и снижением покупательной способности населения. Однако в 2017 и 2019 гг. зафиксирован рост натурального объема медицинского рынка, стимулированный восстановлением способности населения к оплате услуг и увеличением предложения медицинских услуг. К 2019 г. количество медицинских приемов в Самаре достигло 15,07 млн случаев, что на 5,1 % превышает показатель 2015 г. Медицинский рынок Самары демонстрирует активный рост, ежегодно увеличивая количество медицинских учреждений. В регионе открываются как частные, так и государственные медицинские центры различного профиля. Принимаемые в регионе меры, включая реализацию национальных проектов и модернизацию системы здравоохранения, направлены на решение критической проблемы уменьшения численности населения из-за высокой смертности и низкой рождаемости. Основной причиной смертности в Самаре являются заболевания системы кровообращения. В ответ на это в городе и области были открыты новые сосудистые отделения и усовершенствованы существующие медучреждения, что способствовало повышению уровня выявляемости болезней, превышающему среднероссийские показатели. В период с 2020 по 2024 г. прогнозируется положительная динамика в натуральном объеме медицинского рынка Самары, ожидается рост на 6,8 %, достигнув 16,28 млн приемов к 2024 г. Амбулаторно-поликлинические организации в Самаре, как и в других регионах, включают в себя широкий спектр медицинских учреждений, предназначенных для предоставления разнообразных медицинских услуг населению.

Краткий обзор динамического анализа

В исследовании мощности (числа посещений в смену) амбулаторно-поликлинических организаций были применены методы динамического и канонического анализов. Исходные данные для анализа были получены с сайта Федеральной службы государственной статистики (Росстат), что обеспечило доступ к актуальной и объективной информации о числе обращений в медицинские учреждения.

С помощью канонического анализа были исследованы связи между количеством посещений амбулаторно-поликлинических учреждений и различными экологическими факторами, такими как уровень загрязнения воздуха, количество отходов, качество воды: все они влияют на мощность (посещаемость) в медицинских учреждениях. Анализ данных предоставляет возможность установить факт влияния экологических условий на здоровье населения.

Первый этап динамического анализа предполагает изучение изменений величин во времени и включает определение таких показателей, как абсолютный прирост, темп роста и темп прироста. Так, если анализ цепных показателей часто используется для изучения

последовательных изменений величин, то базисный подход классически применяется для сравнения данных на одном уровне или для определения относительного изменения по отношению к начальному уровню [21].

Абсолютный прирост вычисляется по формулам:

$$\Delta_y^{\text{п}} = y_i - y_{i-1}; \quad (1)$$

$$\Delta_y^{\text{б}} = y_i - y_0. \quad (2)$$

Темп роста – отношение уровней ряда динамики, это помогает определить, насколько быстро или медленно изменяется величина за определенный период времени. Темп роста находится по формулам:

$$K_p^{\text{п}} = \frac{y_i}{y_{i-1}}; \quad (3)$$

$$K_p^{\text{б}} = \frac{y_i}{y_0}. \quad (4)$$

Темп прироста позволяет оценить скорость изменения величины независимо от ее масштаба. Этот показатель часто используется для анализа темпа роста в натуральных единицах и для прогнозирования будущих значений. Его можно вычислить по формулам:

$$T_{\text{пр}}^{\text{п}} = \frac{\sum \Delta_{y_{\text{п}}}}{y_{i-1}}; \quad (5)$$

$$T_{\text{пр}}^{\text{б}} = \frac{\sum \Delta_{y_{\text{б}}}}{y_0}. \quad (6)$$

При анализе временного ряда необходимо учитывать, что наблюдаемые изменения могут быть подвержены различным воздействиям, таким как тренды, сезонные колебания и другие. Проверка на стационарность позволяет выявить наличие любых систематических изменений в статистических свойствах ряда, которые могут повлиять на интерпретацию результатов и принятие решений. Таким образом, переход к анализу стационарности дополняет первоначальный анализ изменений величин, обеспечивая более надежную основу для принятия решений на основе временных данных. Значение проверки на стационарность в рамках анализа заключается в том, что стационарные временные ряды обычно легче анализировать и моделировать, поскольку они имеют более предсказуемые характеристики. Более того, многие методы анализа временных рядов, к примеру, модель ARIMA (авторегрессионные интегрированные скользящие средние), требуют стационарности данных для правильного применения.

Для анализа стационарности временных рядов применяется метод, включающий в себя разбиение исходных данных на несколько периодов и последующую проверку гипотез о равенстве дисперсий и средних значений временного ряда. Этот метод предусматривает деление временного ряда на k равных периодов, где значение k , как правило, устанавливается равным «3» для обеспечения достаточного объема данных для анализа.

Для каждого из k периодов рассчитываются средние значения \bar{X}_i и дисперсии S_i^2 временного ряда.

Для анализа равенства средних применяется t -критерий Стьюдента, основная гипотеза которого заключается в том, что средние значения по всем периодам совпадают. Расчет t -статистики осуществляется по формуле

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}, \quad (7)$$

где \bar{X}_1, \bar{X}_2 – средние значения первого и второго периодов; S_1^2, S_2^2 – дисперсии первого и второго периодов; n_1, n_2 – количество наблюдений в каждом периоде.

Превышение рассчитанного значения t -статистики над критическим указывает на отвержение нулевой гипотезы о равенстве средних.

Для проверки гипотезы о равенстве дисперсий используется F -критерий, основанный на предположении об идентичности дисперсий по всем интервалам. F -статистика определяется как отношение дисперсии между группами к дисперсии внутри групп. Подобно t -критерию, превышение рассчитанной F -статистики над ее критическим значением влечет отвержение нулевой гипотезы о равенстве дисперсий [20]. Эти методы позволяют с высокой точностью оценить стационарность временных рядов, выявляя любые статистически значимые изменения в их структуре.

Таким образом, данный метод выступает как мощный инструмент в руках исследователей и аналитиков, стремящихся к точному пониманию динамики и стабильности изучаемых явлений.

После подтверждения или достижения стационарности временного ряда можно также провести анализ автокорреляции. Автокорреляция позволяет выявить взаимосвязи между последовательными значениями временного ряда в различные моменты времени, что особенно важно для идентификации внутренних структурных зависимостей.

В контексте стационарных временных рядов коэффициент автокорреляции помогает в выявлении и моделировании периодических и трендовых компонент, которые могут присутствовать в данных. Это знание используется для настройки и оптимизации прогностических моделей, например, авторегрессионных (AR), интегрированных (I) и моделей скользящего среднего (MA), а также их комбинаций, как в модели ARIMA. Эти модели эффективно учитывают как краткосрочные, так и долгосрочные зависимости, что позволяет исследователям формировать точные и обоснованные прогнозы будущих значений временного ряда.

Коэффициент автокорреляции уровней ряда первого порядка будет иметь следующую формулу:

$$r_1 = \frac{\sum_{t=2}^n (y_t - \bar{y}_1) \cdot (y_{t-1} - \bar{y}_2)}{\sqrt{\sum_{t=2}^n (y_t - \bar{y}_1)^2 \cdot (y_{t-1} - \bar{y}_2)^2}}; \quad (8)$$

$$\bar{y}_1 = \frac{\sum_{t=2}^n y_t}{n-1}; \quad \bar{y}_2 = \frac{\sum_{t=2}^n y_{t-1}}{n-1}. \quad (9)$$

Затем определяется максимальный лаг, который зависит от объема доступных данных и специфики задачи анализа, чтобы учесть как краткосрочные, так и долгосрочные зависимости в данных.

С целью минимизации влияния аномальных значений, которые могут исказить прогнозы, временные ряды обрабатываются методами сглаживания. Применение простой скользящей средней является одним из наиболее активно используемых подходов для этой цели. После проведения сглаживания временных рядов методом скользящей средней важно провести оценку погрешностей, которые могут возникнуть в результате этой процедуры.

После анализа погрешностей, связанных с применением скользящей средней для сглаживания временных рядов, важным этапом является исследование наличия трендовой компоненты в данных. Для этой цели могут использоваться различные статистические критерии, среди которых особое внимание уделяется критерию «восходящих и нисходящих» серий, а также критерию, основанному на медиане выборочной совокупности. После проведения всех подготовительных этапов, сглаживания временных данных и проверки гипотез о наличии тренда, можно приступить к главной составляющей динамического анализа – прогнозированию показателей на основе обработанных данных.

Подбор модели прогнозирования осуществляется с использованием программных инструментов, таких как Microsoft Excel, и основывается на максимизации коэффициента детерминации R^2 , который определяет долю вариативности зависимой переменной, объясняемую моделью. Коэффициент детерминации рассчитывается по формуле

$$R^2 = \frac{\sum(\hat{y}_{xi} - \bar{y})^2}{\sum(y_i - \bar{y})^2}. \quad (10)$$

На основании выбранной модели формируется прогноз будущих значений. Формула прогноза зависит от типа подобранной модели (линейной, полиномиальной, экспоненциальной и т.д.) и включает в себя коэффициенты, полученные в результате статистического анализа.

Для оценки предсказанных значений строится доверительный интервал. Интервальное оценивание используется в статистическом анализе для того, чтобы указать не только оценку неизвестного параметра, но и неопределенность, связанную с этой оценкой. Для построения интервала используют формулу

$$\hat{y}_t - t_\alpha \cdot S_{\hat{y}_t} \leq y_p \leq \hat{y}_t + t_\alpha \cdot S_{\hat{y}_t}, \quad (11)$$

где $S_{\hat{y}_t}$ – стандартная ошибка; t_α – статистика Стьюдента.

Для глубокого понимания динамики временных рядов и их эффективного прогнозирования используются различные модели, среди которых выделяются авторегрессионные модели (AR), модели скользящей средней (MA) и авторегрессионные интегрированные модели скользящей средней (ARIMA). Каждая из этих моделей обладает уникальными характеристиками, которые делают их применимыми к определенным типам временных рядов.

Авторегрессионная модель (AR) предполагает, что текущее значение временного ряда может быть выражено через его предыдущие значения с добавлением случайной ошибки [23]. Модель AR (p) описывается следующим уравнением:

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_p y_{t-p} + \varepsilon_t, \quad (12)$$

где y_t – текущее значение ряда; a_0, a_1, a_2, a_p – параметры модели; p – порядок авторегрессии, указывающий на количество лагов (задержек), используемых в модели; ε_t – случайная ошибка в момент времени t .

Модель скользящей средней (МА) моделирует временной ряд через линейную комбинацию прошедших случайных компонент. Формула модели имеет вид:

$$y_t = a_0 + a_1\varepsilon_{t-1} + a_2\varepsilon_{t-2} + \dots + a_p\varepsilon_t. \quad (13)$$

Модель ARIMA представляет собой мощный инструмент анализа и прогнозирования временных рядов. Она сочетает в себе три ключевых компонента: авторегрессию (AR), интегрирование (I) и скользящую среднюю (МА), что позволяет ей эффективно работать как со стационарными, так и с нестационарными данными.

Авторегрессионный компонент (AR) модели ARIMA описывает текущее значение временного ряда через линейную комбинацию его предыдущих значений. Это позволяет учитывать влияние прошлых наблюдений на будущие в контексте временных зависимостей [24].

Модель скользящей средней (МА) используется для моделирования текущего значения ряда через функцию прошлых ошибок прогноза. Это помогает уловить случайные колебания и выбросы в данных, которые не могут быть объяснены только через авторегрессию.

Эта модель предоставляет аналитикам и научным сотрудникам инструменты для прогнозирования будущих показателей на основании прошлых данных, охватывая как продолжительные тренды и повторяющиеся циклы, так и более кратковременные колебания.

Теоретические аспекты канонического анализа

Канонический анализ открывает новую перспективу, предоставляя возможность изучить взаимосвязи между различными измерениями и показателями работы амбулаторно-поликлинических учреждений. Этот метод позволяет не только углубить понимание динамики, но и выявить скрытые закономерности, которые могут оказать значительное влияние на эффективность и доступность медицинских услуг.

Канонический анализ начинается с определения двух наборов переменных, между которыми исследователь предполагает наличие взаимосвязи [24]. Эти переменные, будучи количественно измеримыми, служат основой для дальнейшего анализа. Затем в центре внимания находится процесс вычисления канонических корреляций, который направлен на нахождение линейных комбинаций переменных в каждом наборе таким образом, чтобы корреляция между этими комбинациями была максимальной. Этот процесс не только выявляет степень связи между наборами, но и позволяет формировать канонические переменные, представляющие собой эти оптимальные комбинации [25].

Согласно каноническому анализу, рассчитываются ковариационные или корреляционные матрицы для каждого набора переменных. Определим матрицы ковариаций между переменными первого набора (X) и второго набора (Y) как S_{XX} и S_{YY} соответственно, а ковариационную матрицу между наборами как S_{XY} . Затем матрица A может быть вычислена как результат произведения обратных матриц ковариаций на матрицу ковариаций между наборами:

$$A = S_{XX}^{-1}S_{XY}S_{YY}^{-1}S_{YX}. \quad (14)$$

Далее для матрицы A находятся собственные векторы (v) и собственные значения (λ^2). Собственные векторы и значения играют ключевую роль, поскольку они используются для формирования канонических переменных. Собственные значения λ определяют степень корреляции между соответствующими каноническими переменными, в то время как собственные векторы v указывают на направления максимальной вариации.

Канонические переменные U и V формируются как линейные комбинации исходных переменных с использованием собственных векторов в качестве коэффициентов. Для каждой пары канонических переменных U_i и V_i , где i указывает на номер пары, формулы выглядят следующим образом:

$$U = a_1x_1 + a_2x_2 + \dots + a_gx_g; \quad (15)$$

$$V = b_1y_1 + b_2y_2 + \dots + b_py_p. \quad (16)$$

Для оценки значимости канонических корреляций используется критерий Бартлетта, который применяется для проверки нулевой гипотезы о том, что все корреляции между двумя наборами переменных равны нулю. Проверить гипотезу можно с помощью χ^2 -критерия:

$$\chi^2 = -\left[n - 1 - \frac{1}{2}(p + q + t) \right] \ln W_0; \quad (17)$$

$$W_0 = 1 - \lambda_i^2. \quad (18)$$

Если значение χ^2 больше табличного значения критерия при выбранном уровне значимости, то можно утверждать, что по крайней мере одна из канонических корреляций статистически значима. Чтобы проверить значимость второго коэффициента канонической корреляции, необходимо рассчитать величину W_1 , которая равна:

$$W_1 = \prod_{i=1}^p (1 - \lambda_i^2). \quad (19)$$

И затем определить χ^2 по формуле

$$\chi^2 = -\left[n - 2 - \frac{1}{2}(p + q + 1) \right] \ln W_1. \quad (20)$$

Комплексный подход к анализу, сочетающий в себе динамический и канонический анализы с последующей проверкой статистической значимости, позволяет получить глубокое понимание динамики и структуры посещаемости амбулаторно-поликлинических учреждений. Полученные результаты могут служить основой для разработки стратегических решений и улучшения работы медицинских учреждений, направленных на повышение доступности и качества медицинской помощи для населения.

Результаты динамического анализа показателей посещаемости амбулаторно-поликлинических организаций в Самарской области

В современных условиях популяризации здорового образа жизни амбулаторно-поликлинические организации играют важную роль в оказании медицинской помощи и поддержании здоровья граждан. Их деятельность напрямую связана с посещаемостью пациентов, что, в свою очередь, влияет на эффективность и качество предоставляемых услуг, поэтому важно анализировать данные показатели. В данном исследовании использована официальная статистика посещаемости амбулаторно-поликлинических организаций, предоставленная Росстатом на период с 2004 до 2021 г. Для наилучшего понимания информация о показателях посещаемости отображается в виде графиков. Данные показателей представлены на рис. 2.

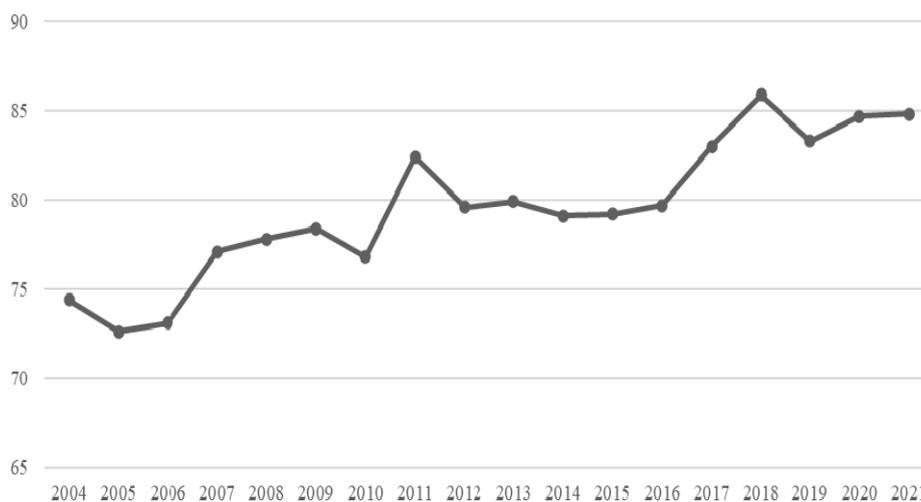


Рис. 2. График посещаемости амбулаторно-поликлинических организаций в Самарской области, тыс. посещений в смену

Изучение графика выявляет мягкую динамику изменения показателя с течением времени, начиная с отметки в 74,4 в 2004 г. и демонстрируя постепенный рост до приблизительно 85,9 к 2019 г. Следует отметить стабилизацию в 2020 и 2021 гг. на уровне около 84,7–84,8, что свидетельствует о некоторой стабильности в последние годы исследования. График будет подниматься и падать в зависимости от изменения показателя, но общий тренд будет восходящим с некоторыми колебаниями. Минимальное значение приходится на начало исследуемого периода (2005), а максимальное значение на конец (2018). Разница между данными значениями составляет 13,3 тысячи, а количество посещений в 2018 г. примерно на 18 % превышает количество посещений в 2005 г.

Для более глубокого понимания тенденции и динамики развития в сфере здравоохранения необходимо рассмотреть количественные характеристики: абсолютный прирост, темп роста, темп прироста, представленные в табл. 1.

Таблица 1

Результаты количественной оценки посещаемости амбулаторно-поликлинических организаций в Самарской области

t	y_t	Абсолютный прирост		Темп роста, %		Темп прироста, %	
		цепной	базисный	цепной	базисный	цепной	базисный
2004	74,4	–	–	–	–	–	–
2005	72,6	1,8	–7	97,58	91,21	–2,42	–8,79
2006	73,1	–0,5	–6,5	100,69	91,83	0,69	–8,17
2007	77,1	–4	–2,5	105,47	96,86	5,47	–3,14
2008	77,8	–0,7	–1,8	100,91	97,74	0,91	–2,26
2009	78,4	–0,6	–1,2	100,77	98,49	0,77	–1,51
2010	76,8	1,6	–2,8	97,96	96,48	–2,04	–3,52
2011	82,4	–5,6	2,8	107,29	103,52	7,29	3,52
2012	79,6	2,8	0	96,60	100,00	–3,40	0,00
2013	79,9	–0,3	0,3	100,38	100,38	0,38	0,38
2014	79,1	0,8	–0,5	99,00	99,37	–1,00	–0,63
2015	79,2	–0,1	–0,4	100,13	99,50	0,13	–0,50
2016	79,7	–0,5	0,1	100,63	100,13	0,63	0,13
2017	83	–3,3	3,4	104,14	104,27	4,14	4,27

t	y_t	Абсолютный прирост		Темп роста, %		Темп прироста, %	
		цепной	базисный	цепной	базисный	цепной	базисный
2018	85,9	-2,9	6,3	103,49	107,91	3,49	7,91
2019	83,3	2,6	3,7	96,97	104,65	-3,03	4,65
2020	84,7	-1,4	5,1	101,68	106,41	1,68	6,41
2021	84,8	-0,1	5,2	100,12	106,53	0,12	6,53

Для расчета средних значений в качестве базисного года был взят 2012 г., который наиболее приближен к значению среднего уровня всего временного ряда. Исходя из представленных данных, можно сказать, что количество посещений амбулаторно-поликлинических организаций в период с 2004 по 2021 г. возрастало в среднем на 612 посещений в год. Средний темп роста составил 100,77 %, что указывает на увеличение числа посещений в среднем примерно на 0,77 % ежегодно. Эти данные свидетельствуют о заметном, хоть и небыстром увеличении обращений за медицинской помощью, выражая тенденцию к постепенному улучшению доступности и популярности амбулаторных услуг.

Для того чтобы убедиться в этом, можно проанализировать диаграмму цепного абсолютного прироста, представленную на рис. 3.

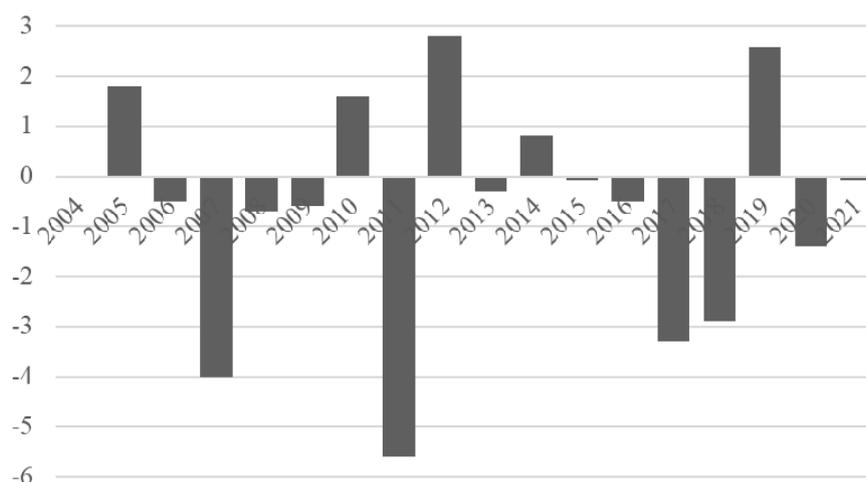


Рис. 3. Диаграмма цепного абсолютного прироста посещаемости амбулаторно-поликлинических организаций в Самарской области

На основе представленного графика можно сделать вывод, что динамика посещений амбулаторно-поликлинических организаций имеет ярко выраженную нестабильность. Далее в работе представлены результаты проверки данных ряда на стационарность. В соответствии с используемой методикой данные были разделены на три периода: 2004–2009 гг. – это первый период, 2010–2015 гг. – второй период, и третий период с 2016 по 2021 г. Для каждого из них были определены статистические показатели, а именно средний уровень, дисперсия и стандартное отклонение. Для первого периода данные показатели соответственно равны 75,57 (среднее значение), 6,32 (дисперсия), 2,51 (стандартное отклонение). Аналогично, для второго периода: 79,53 (среднее значение) 2,69 (дисперсия) 1,64 (стандартное отклонение), и третьего периода: 84,34 (среднее значение) 1,41 (дисперсия) 1,19 (стандартное отклонение). Используя указанные статистические характеристики, были рассчитаны наблюдаемые значения критериев Фишера ($F_{\text{набл}}$) и Стьюдента ($t_{\text{набл}}$) при про-

верке гипотез на стационарность уровней временного ряда. При проверке гипотезы о равенстве дисперсий рассчитанные значения наблюдаемой статистики критерия сопоставлялись с табличными значениями. В результате на уровне значимости $\alpha = 0,05$ попарное сравнение групповых дисперсий показало их равенство, то есть в каждом случае табличное значение $F_{\text{крит}}$ превышает вычисленное $F_{\text{набл}}$. Данный результат указывает на однородность данных и одинаковую меру рассеяния. Однако при тестировании гипотезы о равенстве средних во всех случаях найденное значение $t_{\text{набл}}$ больше табличного значения $t_{\text{крит}}$. В результате была принята конкурирующая гипотеза H_1 (средние значения статистически значимо различаются), что указывает на нестационарность ряда. Следовательно, данные требуют проведения процедуры сглаживания. В работе также были рассчитаны коэффициенты автокорреляции для лагов от одного до пяти. Перечислим эти значения. Для лага 1 коэффициент автокорреляции составил 0,82, для лага 2 – 0,732 для лага 3 – 0,646, для лагов 4 и 5 – 0,635 и 0,529 соответственно. Наличие высоких значений автокорреляции указывает на сильную связь между последовательными значениями временного ряда, что может указывать на сезонные или циклические паттерны. Прогнозирование изменений посещаемости медицинских учреждений имеет большое значение для планирования ресурсов, управления персоналом, оптимизации расходов и обеспечения качественной и доступной медицинской помощи. Прогнозирование изменений показателей посещаемости медицинских учреждений возможно с использованием представления данных временных рядов в виде трендов. Этот процесс начинается с проведения сглаживания данных. В рамках данного исследования для выравнивания данных временного ряда был взят метод простой скользящей средней. Результаты сглаживания для трех интервалов представлены с учетом восстановления утраченных при сглаживании данных в табл. 2.

Таблица 2

Результаты сглаживания данных, тыс. посещений в смену

t	y_t	Интервал сглаживания		
		$L = 3$	$L = 5$	$L = 7$
2004	74,4	72,05	73	73,85
2005	72,6	73,367	74	74,481
2006	73,1	74,267	75	75,112
2007	77,1	76	75,8	75,743
2008	77,8	77,767	76,64	76,886
2009	78,4	77,667	78,5	77,886
2010	76,8	79,2	79	78,857
2011	82,4	79,6	79,42	79,143
2012	79,6	80,633	79,56	79,343
2013	79,9	79,533	80,04	79,529
2014	79,1	79,4	79,5	80,414
2015	79,2	79,333	80,18	80,914
2016	79,7	80,633	81,38	81,443
2017	83	82,867	82,22	82,129
2018	85,9	84,0667	83,32	82,9429
2019	83,3	84,6333	84,34	83,5762
2020	84,7	84,2667	85,38	84,2095
2021	84,8	84,3667	86,42	84,8429

Графическая иллюстрация для интервала сглаживания $L = 7$ представлена на рис. 4.

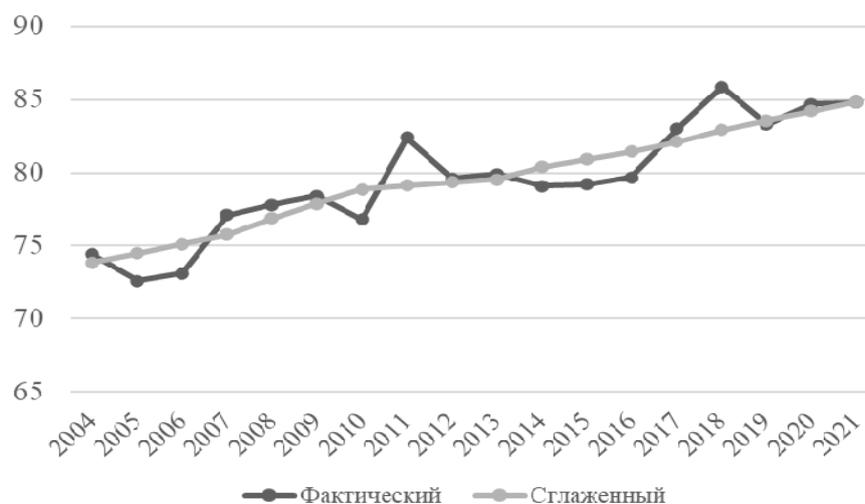


Рис. 4. График показателей посещаемости для периода сглаживания $L = 7$

На каждом из периодов прослеживается возрастающий линейный тренд. Это может означать, что данные имеют тенденцию к увеличению с постоянной скоростью. Возможно, имеется какая-то основная причина, которая вызывает этот тренд, например, увеличение численности населения, изменение экологии региона, улучшение доступности медицинской помощи, повышение осведомленности населения о профилактике заболеваний и важности регулярных медицинских осмотров, а также изменение в рабочем графике и образе жизни людей. Для того чтобы выбрать период сглаживания, необходимо применить оценку погрешности сглаживания верхних и нижних восстановленных значений, которая позволяет определить, насколько точно сглаженный ряд представляет исходные данные. Численное значение погрешности восстановленных данных не превосходит 0,032. Наименьшую погрешность при восстановлении утраченных данных имеют уровни ряда для интервала сглаживания $L = 7$. Также это иллюстрируется путем сравнения значений среднего абсолютного прироста сглаженных функций с исходным (табл. 3).

Таблица 3

Результаты сравнения среднего абсолютного прироста для исходных и сглаженных данных

Средний абсолютный прирост на 1000 посещений			
Исходные данные	$L = 3$	$L = 5$	$L = 7$
0,612	0,725	0,79	0,647

Как видно из данных табл. 3, минимальное расхождение между исходным средним абсолютным приростом и рассчитанным по сглаженным данным наблюдается для периода сглаживания $L = 7$. Используя медианный критерий и критерий «восходящих и нисходящих» серий, проводилась проверка гипотезы о наличии тренда. По результатам проверки установлено, что подтверждается гипотеза о наличии тренда по медианному критерию и критерию «восходящих и нисходящих» серий.

Модели прогнозирования

Следующим этапом исследования является определение трендовой модели. В работе рассматривались как линейные, так и нелинейные модели, был сделан вывод, что исследуемые данные наиболее точно описывает полиномиальная функция третьей степени (рис. 5).

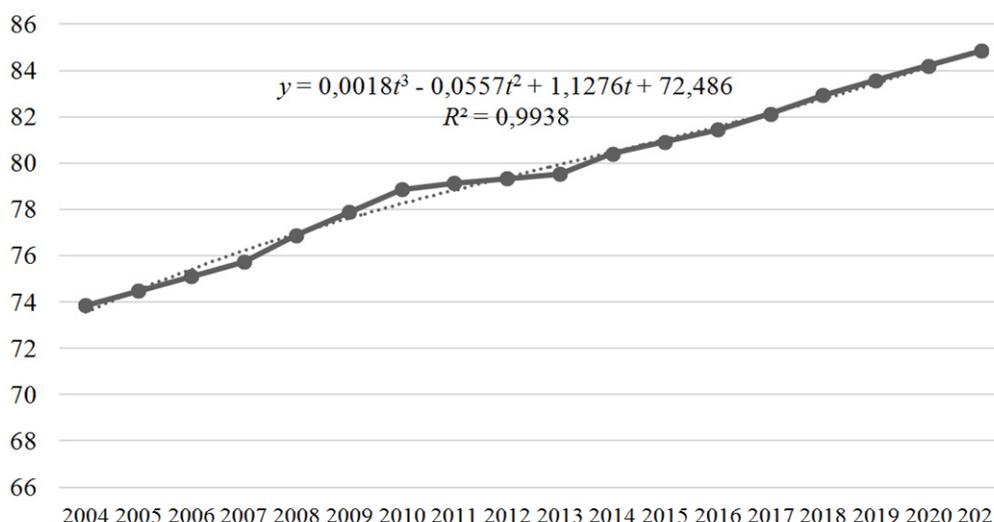


Рис. 5. График тренда посещаемости амбулаторно-поликлинических организаций в Самарской области

Таким образом, функция для тренда имеет следующий вид:

$$y(t) = 0,0018t^3 - 0,0557t^2 + 1,127t + 72,486.$$

Используя данную функцию, можно провести оценку прогноза численности посещаемых амбулаторно-поликлинических учреждений. Это позволяет заранее подготовиться к возможным изменениям и адаптировать работу организаций к ожидаемым колебаниям посещаемости. При помощи данной и других функции тренда были получены следующие прогнозные значения на три года, представленные в табл. 4.

Таблица 4

Результаты прогноза числа посещений амбулаторно-поликлинических учреждений

Год	Линейная модель, тыс. посещений	Параболическая модель, тыс. посещений	Полиномиальная модель, тыс. посещений
2022	85,482	85,116	86,149
2023	86,11	85,628	87,158
2024	86,738	86,128	88,272

Исходя из результатов, прогнозные значения посещаемости амбулаторно-поликлинических организаций близки к реальным значениям, и любая из моделей прогнозирования показывает дальнейший умеренный рост рассматриваемых показателей в рамках периода упреждения, равного трем годам. Интервальная оценка в трендовых моделях имеет большое значение, поскольку она позволяет оценить стабильность и надежность прогнозов. Это позволяет принимать обоснованные решения на основе прогнозов и уменьшить риски ошибочных действий. Доверительные интервалы прогнозных уровней для каждой из видов трендовых моделей – линейной, параболической и полиномиальной – представлены в табл. 5.

Интервальные оценки позволяют определить степень уверенности в прогнозах по каждой модели. Чем меньше разница между верхней и нижней оценкой, тем точнее прогноз.

Доверительные интервалы трендовых моделей

Модель	Год					
	2022		2023		2024	
	Граница		Граница		Граница	
	левая	правая	левая	правая	левая	правая
Линейная	78,37	92,595	80,64	89,592	81,673	90,625
Параболическая	78,998	93,223	81,152	90,103	82,682	91,634
Полиномиальная	79,625	93,85	81,653	90,604	83,796	92,747

Далее в работе представлены результаты по следующим моделям: модели AR (авторегрессионная модель), MA (модель скользящего среднего), ARIMA (авторегрессионная интегрированная модель скользящего среднего). Они являются важными инструментами для анализа рядов динамики и прогнозирования их значений. В модели AR используется линейная комбинация предыдущих значений временного ряда с коэффициентами, которые определяются в процессе обучения модели. С помощью пакета анализа Excel проводится расчет параметров моделей и анализируется регрессионная статистика. В результате получены следующие модели AR, описываемые формулами (21)–(24).

$$y_t = 13,975 + 0,81y_{t-1}; \quad (21)$$

$$y_t = 10,689 + 0,572y_{t-1} + 0,283y_{t-2}; \quad (22)$$

$$y_t = 6,862 + 0,52y_{t-1} + 0,21y_{t-2} + 0,169y_{t-3}; \quad (23)$$

$$y_t = 13,654 + 0,499y_{t-1} + 0,046y_{t-2} + 0,04y_{t-3} + 0,305y_{t-4}. \quad (24)$$

Из этих уравнений можно сделать общий вывод, что количество посещений амбулаторно-поликлинических организаций увеличивается с течением времени, и это увеличение может быть спрогнозировано на основе предыдущих значений посещений. Коэффициенты перед прошлыми значениями посещений показывают, насколько сильно предыдущие значения влияют на текущее количество посещений. Таким образом, можно сделать вывод, что тренд посещений амбулаторно-поликлинических организаций в данной модели является положительным и может быть использован для прогноза будущих значений посещений. Как видно из всех уравнений, доля посещения убывает с ростом лага модели: для модели второго порядка – примерно в два раза; для модели четвертого порядка уже примерно в 10 раз, но при этом сохраняется положительная тенденция. Однако накопительная величина посещений имеет тенденцию к росту.

Следующая MA – модель, которая оценивает текущее значение временного ряда на основе предыдущих случайных компонент и шумов. В модели MA используется линейная комбинация предыдущих остатков модели с коэффициентами, которые также определяются с помощью пакета MS Excel. Полученные в работе модели скользящего среднего представлены соотношениями (25)–(28).

$$y_t = 79,236 + E_{t-1}; \quad (25)$$

$$y_t = 78,883 - 0,124E_{t-1} + 1,124E_{t-2}; \quad (26)$$

$$y_t = 78,594 - 0,13E_{t-1} + 0,592E_{t-2} + 0,538E_{t-3}; \quad (27)$$

$$y_t = 77,686 - 0,868E_{t-1} + 0,262E_{t-2} + 0,525E_{t-3} + 2,13E_{t-4}. \quad (28)$$

Как видно из представленных уравнений МА по посещениям амбулаторно-поликлинических организаций, начальный уровень посещаемости во всех моделях примерно одинаковый. Случайности ближайшего периода для моделей второго, третьего и четвертого порядков оказывают снижение на уровень посещений амбулаторно-поликлинических организаций. Практически одинаковый уровень снижения посещений в моделях второго и третьего порядков и уже в 7 раз возрастает снижение в модели четвертого порядка по сравнению с моделями второго порядка. Лаги, начиная со второго и выше, демонстрируют положительную компенсацию, т.е. приводят к росту уровня числа посещений амбулаторно-поликлинических организаций. В качестве случайностей можно рассматривать дополнительные факторы, такие как сезонность, праздники, эпидемии и другие внешние воздействия.

Следующая модель ARIMA – это модель, которая объединяет в себе авторегрессионные и скользящие средние компоненты, а также возможность интегрирования временного ряда для удаления тренда. Согласно регрессионной статистике, среди лагов и остатков значимыми являются лаги и остатки первого порядка, поэтому модель и строилась с учетом этого обстоятельства.

Общий вид модели ARIMA выражается соотношением

$$y_t = 13,975 + 0,817y_{t-1} + E_{t-1}.$$

Из представленного уравнения ARIMA для посещений амбулаторно-поликлинических организаций следует вывод:

1) уровень посещений в текущем периоде имеет положительный коэффициент пропорциональности с уровнем предшествующего периода, следовательно, приводит к росту числа посещений амбулаторно-поликлинических учреждений;

2) на уровень посещений также влияет случайное возмущение из предыдущего периода, что означает возможность влияния различных внешних факторов, не учтенных в модели. Это обстоятельство предполагает выявления влияния факторов, которые могут быть как случайными, так и не случайными. Для этой цели проводится канонический анализ групп медицинских и экологических показателей.

Канонический анализ медицинских и экологических показателей

Следующим шагом в исследовании является изучение влияния внешних факторов на эти изменения. Экологические условия, охватывающие широкий спектр факторов, от качества воздуха до доступности зеленых зон, играют значительную роль в определении общего состояния здоровья населения, что, в свою очередь, отражается на посещаемости медицинских учреждений. Такой подход не только расширяет наше понимание влияния окружающей среды на здоровье населения, но и предоставляет нам инструменты для прогнозирования будущих изменений в потребностях в медицинских услугах, основываясь на предполагаемых экологических сценариях. Для исследования были использованы данные, взятые с официального сайта Росстата.

Первым этапом анализа выдвигается основная гипотеза – экологические факторы оказывают влияние на посещаемость медицинских учреждений и заболеваемость населения.

Для изучения связей между экологическими факторами (X) и здоровьем населения (Y) были взяты динамические данные для всех регионов РФ. Признаки X_1 , X_2 и X_3 относятся к экологическим факторам, которые могут влиять на здоровье населения. Они выбраны для канонического анализа, потому что позволяют исследовать связь между уровнем экологической безопасности регионов и состоянием здравоохранения в них. Y_1 и Y_2 относятся к показателям в сфере здравоохранения. Их анализ может выявить, как экологические условия влияют на потребность в медицинских услугах и общее состояние здоровья населения. Подробная расшифровка используемых факторов представлена в табл. 6.

Таблица 6

Факторы канонического анализа

Показатель	Расшифровка
X_1	Выбросы загрязняющих атмосферу веществ стационарными и передвижными источниками
X_2	Поступление загрязняющих веществ со сточными водами в водоемы по Российской Федерации
X_3	Образование отходов производства и потребления
Y_1	Мощность (число посещений в смену) амбулаторно-поликлинических организаций, на конец года
Y_2	Заболеваемость населения по основным классам болезней

Исходные данные канонического анализа представлены в табл. 7.

Таблица 7

Исходные данные канонического анализа

№ п/п	Год	X			Y	
		X_1	X_2	X_3	Y_1	Y_2
		тыс. тонн	млрд м ³	млн тонн	тыс. посещений	тыс.
1	2004	35751,0	51,3	2644,3	3577,5	106287,0
2	2005	35835,0	50,9	3035,5	3637,9	105886,0
3	2006	35510,0	51,4	3519,4	3646,2	108842,0
4	2007	35532,0	51,4	3899,3	3673,9	109571,0
5	2008	33952,0	52,1	3876,9	3651,0	109590,0
6	2009	32754,0	47,7	3505,0	3657,2	113877,0
7	2010	32353,0	49,2	3734,7	3685,1	111428,0
8	2011	32628,0	48,1	4303,3	3727,7	113921,8
9	2012	32469,0	45,5	5007,9	3780,4	113688,4
10	2013	32063,0	42,9	5152,8	3799,4	114721,1
11	2014	31228,0	43,7	5168,3	3858,5	114989,0
12	2015	31269,0	42,9	5060,2	3861,0	113926,9
13	2016	31617,0	42,9	5441,3	3914,2	115187,3
14	2017	32068,0	42,6	6220,6	3966,7	114382,2
15	2018	32327,0	40,1	7266,1	3997,8	114840,8
16	2019	22735,0	37,7	7750,9	4072,4	114512,2
17	2020	22228,0	34,3	6955,7	4147,3	111294,3
18	2021	22300,0	35,6	8448,6	4253,4	125022,4

Матрица корреляций для анализа связи между показателями представлена в табл. 8.

Таблица 8

Матрица корреляций

	X_1	X_2	X_3	Y_1	Y_2
X_1	1,000	0,891	-0,845	-0,889	-0,636
X_2	0,891	1,000	-0,934	-0,964	-0,706
X_3	-0,845	-0,934	1,000	0,973	0,764
Y_1	-0,889	-0,964	0,973	1,000	0,745
Y_2	-0,636	-0,706	0,764	0,745	1,000

Наблюдаемые высокие положительные и отрицательные значения корреляционных коэффициентов указывают на сильные линейные зависимости между исследуемыми переменными.

Следующим шагом необходимо построить матрицу A , для которой вычисляются собственные значения λ (характеристические корни) и собственные векторы v . Используя соотношение (14), была получена матрица A :

$$\begin{pmatrix} 0,949 & 0,685 \\ 0,034 & 0,075 \end{pmatrix}. \quad (29)$$

Рассчитанные собственные значения имеют следующие значения $\lambda_1 = 0,975$ и $\lambda_2 = 0,049$. Расчетные значения коэффициентов корреляции соответственно равны 0,987 и 0,221.

Собственные векторы имеют вид:

$$v_1 = \begin{pmatrix} 26,467 \\ 1 \end{pmatrix}; v_2 = \begin{pmatrix} -0,761 \\ 1 \end{pmatrix}.$$

Исходя из полученных собственных значений, согласно соотношениям (15) и (16), были сформированы канонические переменные:

$$\begin{cases} U_1 = -3,06X_1 - 9,13X_2 + 16,10X_3; \\ V_1 = 26,467Y_1 + Y_2 \end{cases}; \quad (30)$$

$$\begin{cases} U_2 = 2,114X_1 + 6,211X_2 + 8,068X_3; \\ V_1 = -0,761Y_1 + Y_2 \end{cases}. \quad (31)$$

Линейные комбинации групп показателей помогают определить структуру, которая лучше всего описывает взаимосвязь между наборами, позволяя анализировать сложные многомерные связи более простым и интерпретируемым образом. На основе представленных канонических переменных можно сделать вывод, что первый набор канонических переменных (U_1 и V_1) предполагает более сильную линейную связь между переменными X и Y , чем второй набор (U_2 и V_2). Коэффициенты при переменных X в U_1 и при Y в V_1 значительно выше по абсолютной величине, что может указывать на большую вариативность этих переменных, а значит, и на более высокую объясняющую способность модели для этой пары канонических переменных.

Основываясь на значениях коэффициентов канонических переменных и их связи с экологическими и медицинскими факторами, можно предположить, что каноническая переменная, охватывающая линейной комбинацией факторы экологии, подвержена сильному снижению при ухудшении экологических условий (выбросы в атмосферу, сточные воды и отходы). Каноническая переменная, связывающая линейной комбинацией показатели мощность амбулаторных организаций и заболеваемость, также имеет значительный рост. В частности, высокие значения коэффициентов при X_3 в U_1 и положительное значение при Y_1 в V_1 могут указывать на то, что возросшая нагрузка на амбулаторные службы связана с увеличением объема отходов. Полученные значения систем канонических переменных позволяют разработать интегральную шкалу для структуризации объектов по уровню медицинской и экологической составляющей. Если ввести пороговую величину для канонических переменных, то это позволит быстро реагировать на изменение ситуаций, характеризующих состояние региона, страны или конкретной области.

Очевидно, улучшение экологической ситуации окажет положительный эффект на общее здоровье населения и снизит нагрузку на медицинские учреждения. Таким образом, стратегии по охране окружающей среды могут быть рассмотрены как важная часть программ общественного здравоохранения.

Далее представлены результаты проверки на статистическую значимость канонических корреляций по критерию Бартлетта (17)–(21). При уровне значимости $\alpha = 0,05$ были получены следующие результаты (табл. 9).

Таблица 9

Проверка на статистическую значимость по критерию Бартлетта

Первый коэффициент корреляции 0,987		Второй коэффициент корреляции 0,221	
W_0	0,03	W_1	0,02
$\ln W_0$	-3,69	$\ln W_1$	-3,739
$\chi_{\text{набл}}^2$	47,96	$\chi_{\text{набл}}^2$	44,869
$\chi_{\text{крит}}^2$	12,59	$\chi_{\text{крит}}^2$	5,991

Как видно, в каждом из случаев $\chi_{\text{набл}}^2$ больше, чем $\chi_{\text{крит}}^2$. Из этого можно сделать вывод, что оба коэффициента статистически значимы. Это свидетельствует о значимом статистическом взаимодействии между выбранными наборами переменных. Обобщая, следует отметить, что как в случае сильной степени тесноты связи между группами показателей двух сфер, так и случае слабой степени тесноты связи, параметры канонических переменных следует рассматривать как границы влияния отдельного показателя на величину суммарного значения всех показателей вместе.

Таким образом, проведенный канонический анализ может быть использован как инструмент прогнозирования уровня влияния групп показателей. Дополнительно канонический анализ способствует выявлению других групп факторов, которые в моделях MA и ARIMA изначально участвуют как неучтенные факторы. Эти параметры могут служить катализатором для дальнейших исследований и практических действий в области здравоохранения и экологии.

Заключение

Представлены результаты оценки прогнозного числа посещаемости амбулаторно-поликлинических организаций в Самарской области. Проведен динамический анализ с прогнозированием посещаемости амбулаторно-поликлинических учреждений и канонический анализ для выявления связи экологических факторов и факторов в сфере здравоохранения. Исследование динамики посещаемости амбулаторно-поликлинических учреждений в Самарской области подчеркнуло их важность в общей системе здравоохранения. Установлено, что как в случае слабой степени тесноты взаимосвязи между группами показателей экологии и групп медицинских факторов, так и в случае сильной степени связи канонические переменные позволяют определять тенденцию влияния каждого отдельного показателя на групповой интегральный фактор. Анализ связей между экологическими факторами и посещаемостью позволил не только глубже понять существующие тенденции. Полученные данные и разработанные на их основе прогнозные модели могут быть базой для разработки стратегий управления, направленных на оптимизацию ресурсов и повышение доступности медицинских услуг для населения.

Список литературы

1. Якимова, Л.А. Исследование качества жизни сельских жителей Красноярского края методом сравнительной социальной интегральной оценки / Л.А. Якимова, А.В. Стрельцова // Наука Красноярья. – 2020. – Т. 9, № 4. – С. 255–267. – DOI 10.12731/2070-7568-2020-4-255-267
2. Кожевников, А.А. Применение индекса наличия проблемы при сравнительном анализе заболеваемости населения Российской Федерации, Сибирского федерального округа и Кемеровской области / А.А. Кожевников, Д.Г. Данцигер, К.В. Часовников // Прикаспийский журнал: управление и высокие технологии. – 2019. – № 2 (46). – С. 22–30.
3. Обзор способов моделирования влияния демографических факторов на параметры социально-экономического развития и систему здравоохранения / А.В. Леонидов, О.В. Обухова, И.П. Шибалков, Д.С. Тюфилин, Н.Г. Куракова, В.И. Климко, В.В. Кулаков // Социальные аспекты здоровья населения. – 2023. – Т. 69, № 4. – DOI 10.21045/2071-5021-2023-69-4-19. – EDN HSNDDDB.
4. Тарабукина, С.М. Основные аспекты направлений развития лекарственного обеспечения в Арктике / С.М. Тарабукина, Н.Б. Дремова // Современные проблемы здравоохранения и медицинской статистики. – 2020. – № 4. – С. 57–70. – DOI 10.24411/2312-2935-2020-00097. – EDN QXDGIS.
5. Яшина, Н.Г. Методика оценки эффективности управления бюджетными ресурсами в сфере здравоохранения / Н.Г. Яшина, И.А. Гришунина, К.С. Яшин // Экономический анализ: теория и практика. – 2015. – № 21 (420). – С. 15–24.
6. Болгарев, Д.В. Моделирование деятельности медицинских организаций на эвакуационном направлении / Д.В. Болгарев, Д.Н. Борисов // Известия Российской военно-медицинской академии. – 2021. – Т. 40, № 1. – С. 59–65. – DOI 10.17816/rmmar64484
7. Латуха, О.А. Модель управления устойчивым развитием медицинской организации, оказывающей помощь в амбулаторных условиях: дис. ... д-ра мед. наук / Латуха Ольга Александровна. – Новосибирск, 2024. – 469 с.
8. Мухачева, А. В. Математическое моделирование качества жизни населения региона / А. В. Мухачева, Е. Я. Пастухова, А. Н. Кирюхина // Вестник Омского университета. Серия:

Экономика. – 2020. – Т. 18, № 1. – С. 149–161. – DOI 10.24147/1812-3988.2020.18(1).149-161. – EDN WUUXOF.

9. Аксенова Е.И., Бессчетнова О.В. Показатели доступности и качества медицинской помощи, обеспечивающие удовлетворенность населения медицинской помощью в различных странах мира: экспертный обзор. – М.: ГБУ «НИИОЗММ ДЗМ», 2021. – 40 с.

10. Набережная, И. Б. Инновационные технологии и их место в борьбе с сердечно-сосудистыми заболеваниями / И. Б. Набережная // Современные проблемы здравоохранения и медицинской статистики. – 2021. – № 4. – С. 457–470. – DOI 10.24412/2312-2935-2021-4-457-470. – EDN OABLYX.

11. Саадалов, Т.Ы. Методика расчета коэффициента корреляции Фехнера и Пирсона, и их области применения / Т.Ы. Саадалов, Р.М. Мырзаibraимов, Ж.Д. Абдуллаева // Бюллетень науки и практики. – 2021. – Т. 7, № 10. – С. 270-276. – DOI 10.33619/2414-2948/71/31. – EDN GNMYZT.

12. Набережная, И.Б. Региональные особенности динамики смертности и объемов оказания высокотехнологичной медицинской помощи / И. Б. Набережная // Российский медицинский журнал. – 2020. – № 26, № 5. – С. 266–273. – DOI 10.17816/0869-2106-2020-26-5-266-273. – EDN YCUUHK.

13. Окрепилов, В.В. Формирование подходов к применению стандартизации для оценки качества жизни / В. В. Окрепилов, И. В. Чудиновских // Петербургский экономический журнал. – 2018. – № 1. – С. 6–15.

14. Макаров, В. Л. Принципы мониторинга качества жизни на основе агент-ориентированных моделей / В. Л. Макаров, В. В. Окрепилов // Вестник Российской академии наук. – 2016. – Т. 86, № 8. – С. 711-718. – DOI 10.7868/S0869587316080089. – EDN WHYYFR.

15. Тумусов Ф.С., Косенков Д.А. Современные тенденции в системе здравоохранения Российской Федерации. – М.: Издание Государственной Думы, 2019. – 80 с.

16. Анализ рынка медицинских услуг в Самаре в 2015–2019 гг., прогноз на 2020–2024 гг. [демоверсия], [Электронный ресурс]. – URL: https://businessstat.ru/images/demo/medicine_samara_demo_businessstat.pdf (дата обращения: 12.11.2024).

17. Грицай, А.А. Авторегрессия (AR, autoregression) [Электронный ресурс]. – URL: <https://fnow.ru/algorithm-comparison/avtoregressia> (дата обращения: 12.11.2024).

18. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы: учебник. – М.: Финансы и статистика, 2000. – 352 с.

19. Федеральная служба государственной статистики [Электронный ресурс]. – URL: <https://rosstat.gov.ru/statistic> (дата обращения: 12.11.2024).

20. Самарастат. Орган Федеральной службы государственной статистики по Самарской области [Электронный ресурс]. – URL: <https://63.rosstat.gov.ru/> (дата обращения: 12.11.2024).

21. Компьютерные технологии анализа данных в эконометрике: монография / Д. М. Дайинбегов. – 3-е изд., доп. – М.: Вузовский учебник: ИНФРА-М, 2018. – 587 с.

22. Прикладная статистика. Основы эконометрики: учеб. для студентов экон. специальностей вузов: в 2 т. / С.А. Айвазян, В.С. Мхитарян. Т. 2. – М.: ЮНИТИ-ДАНА, 2001. – 432 с.

23. Кремер Н.Ш., Путко Б.А. Эконометрика: учебник для вузов / под ред. проф. Н. Ш. Кремера. – М.: ЮНИТИ-ДАНА, 2002. – 311 с.

24. Многомерный статистический анализ в экономике / Л.А. Сошникова, В.Н. Тамашевич, Г. Уебе, М. Шефер. – М.: Юнити-Дана, 1999. – 598 с. – EDN EOFQGD.

References

1. Yakimova L.A., Streltsova A.V. The life quality study of rural population in the Krasnoyarsk region by the method of comparative social integral assessment. *Krasnoyarsk Science*, 2020, vol. 9, no. 4, pp. 255–267. – DOI 10.12731/2070-7568-2020-4-255-267
2. Kozhevnikov A.A., Dantsiger D.G., Chasovnikov K.V. Application of the index of existence of the problem in the comparative analysis of incidence of the population of Russian Federation, Siberian Federal District and the Kemerovo Region. *CASPIAN JOURNAL: Control and High Technologies*, 2019, no. 2 (46), pp. 22–30.
3. Leonidov A.V., Obukhova O.V., Shibalkov I.P., Tufilin D.S., Kurakova N.G., Klimko V.I., Timonin S.A. Review of the methods for modeling the impact of demographic factors on the parameters of socioeconomic development and health care system. *Social aspects of Population Health*, 2023, vol. 69, no. 4, DOI 10.21045/2071-5021-2023-69-4-19
4. Tarabukina S.M., Dremova N.B. Use of the results of the expert survey in the development of directions for the development of drugs in the Arctic. *Current problems of health care and medical statistics*, 2020, no. 4, pp. 57–70.
5. Yashina N.G., Grishunina I.A., Yashin K.S. Methods for assessing the effectiveness of budget resource management in healthcare. *ECONOMIC ANALYSIS: theory and practice*, 2015, no. 21 (420), pp. 15-24.
6. Bolgarev D.V., Borisov D.N. Modeling the activities of medical organizations in the evacuation direction. *Russian Military Medical Academy Reports*, 2021, vol. 40, no. 1. pp. 59-65. – DOI 10.17816/rmmar64484
7. Latukha O.A. A model for managing sustainable development of a medical organisation providing care in outpatient settings. Doctor's Degree dissertation, Novosibirsk, 2024, 469 p.
8. Mukhacheva A.V., Pastukhova E.Ia., Kiryukhina A.N. Mathematical modeling of the quality of life of the region's population. *Herald of Omsk University. Series "Economics"*, 2020, vol. 18, no. 1, pp. 149-161. – DOI 10.24147/1812-3988.2020.18(1).149-161.
9. Aksenova E.I., Besschetnova O.V. Indicators of accessibility and quality of medical care ensuring satisfaction of the population with medical care in different countries of the world: an expert review. Moscow, State Budgetary Institution «Research Institute for Healthcare Organization and Medical Management of Moscow Healthcare Department », 2021, 40 p.
10. Naberezhnaya I.B. Innovative technologies and their seat in struggle against cardiovascular diseases. *Current problems of health care and medical statistics*, 2021, no. 4, pp. 457-470. – DOI 10.24412/2312-2935-2021-4-457-470.
11. Saadalov T., Myrzaibraimov R., Abdullaeva Zh. Calculating procedure for the correlation coefficient of Fechner and Pearson and their application areas. *Bulletin of Science and Practice*, 2021 vol. 7, no. 10, pp. 270-276. – DOI 10.33619/2414-2948/71/31.
12. Naberezhnaia I.B. Mortality dynamics and scope of high technology medical care availability: regional specificities. *Rossiiskii meditsinskii zhurnal (Medical Journal of the Russian Federation, Russian journal)*, 2020, vol. 26, no. 5, pp. 266–273. – DOI 10.17816/0869-2106-2020-26-5-266-273.
13. Okrepilov V.V., Chudinovskikh I.V. Approaches formation to application of standardization for assessment of quality of life. *Peterburgskii ekonomicheskii zhurnal*, 2018, no. 1, pp. 6–15.
14. Makarov, V.L., Okrepilov V.V. Principles of life quality monitoring based on agent-oriented models. *Herald of the Russian Academy of Sciences.*, 2016, vol. 86, no. 8, pp. 322–328. – DOI 10.1134/S1019331616040092

15. Tumusov F.S., Kosenkov D.A. Current trends in the health care system of the Russian Federation. Moscow, Publ. of the State Duma of the Federal Assembly of the Russian Federation, 2019, 80 p.
16. Analysis of the Samara medical services market in 2015-2019, forecast for 2020-2024, available at: https://businessstat.ru/images/demo/medicine_samara_demo_businessstat.pdf (Accessed: 12.11.2024)
17. Gritsai A.A. Autoregression (AR, autoregression), available at: <https://fnow.ru/algorithm-comparison/avtoregressia> (Accessed: 12 November 2024).
18. Dubrov A.M., Mkhitarian V.S., Troshin L.I. Multivariate statistical methods. Moscow, Finansy i statistika, 2000, 352 p.
19. Federal State Statistics Service, available at: <https://rosstat.gov.ru/statistic> (Accessed: 12 November 2024).
20. Samarastat. Federal State Statistical Service of Samara region, available at: <https://63.rosstat.gov.ru/> (Accessed: 12 November 2024).
21. Dayitbegov D.M. Computer technologies for data analysis in econometrics. Moscow, Vuzovskii uchebnyk: INFRA-M, 2018, 587 p.
22. Aivazyan S.A., Mhitaryan V.S. Essentials of econometrics. 2 Parts. Part 2. Moscow, IuNITI-DANA, 2001, 432 p.
23. Kremer N.Sh., Putko B.A. Econometrics. Moscow, IuNITI-DANA, 2002, 311 p.
24. Soshnikova L.A., Tamashevich V.N., Uebe G., Sheffer M. Multivariate statistical analysis in economics. Moscow, IuNITI-DANA, 1999, 598 p.