

УДК 519.218.28

DOI: [10.26102/2310-6018/2025.48.1.024](https://doi.org/10.26102/2310-6018/2025.48.1.024)

Аппроксимации госпитальной статистики выздоровлений от COVID-19

А.В. Боровский¹, А.Л. Галкин², С.С. Дорошенко³

¹Байкальский государственный университет, Иркутск, Российская Федерация

²Институт общей физики им. А.М. Прохорова РАН, Москва, Российская Федерация

³Иркутская ордена «Знак почета» областная клиническая больница, Иркутск, Российская Федерация

Резюме. Госпитальная статистика выздоровлений от COVID-19 по городу Иркутску представлена в виде скорости выздоровления за определенное количество суток от полной группы заболевших. Время выздоровления изменяется от 1 до 182 суток. Количество рассмотренных случаев достигает ~100000 случаев. Для удобства использования данных предложено аппроксимировать таблицу для скорости выздоровления различными видами нелинейных функций. Изучены следующие варианты аппроксимирующих функций: Гауссова, Лоренцева, модифицированная Лоренцева, функция Вейбулла, функции Джонсона. Для сравнения со статистикой применялись методы минимизации среднеквадратичных отклонений аппроксимирующих функций от экспериментальных данных. Использованы метод наименьших квадратов для функций с двумя и тремя параметрами, метод спуска по координатам и метод спуска по градиенту для функций с четырьмя подгоночными параметрами. Показано, что наилучшие результаты подгонки обеспечивает модифицированная Лоренцева функция; с четырьмя параметрами. По степени расхождения с экспериментальной статистикой аппроксимирующие функции расположены в следующем порядке: функция Вейбулла обеспечивает наименее точную подгонку (16,15 %), далее располагается функция Джонсона S_U (10,65 %), несколько лучше подгонка для функции Джонсона S_B (8,49 %), для Гауссовой функции (5,8 %), для Лоренцевой функции подгонка составляет (3,2828 %), наилучшую подгонку дает модифицированная Лоренцева функция (3,2804 %) при определенных приближениях.

Ключевые слова: теория эпидемий, методы оптимизации, спуск по координатам, спуск по градиенту, метод наименьших квадратов, аппроксимация Гаусса, аппроксимация Лоренца, аппроксимация Вейбулла, аппроксимация Джонсона, модифицированное распределение Лоренца.

Для цитирования: Боровский А.В., Галкин А.Л., Дорошенко С.С. Аппроксимации госпитальной статистики выздоровлений от COVID-19. *Моделирование, оптимизация и информационные технологии*. 2025;13(1). URL: <https://moitvivr.ru/ru/journal/pdf?id=1812> DOI: 10.26102/2310-6018/2025.48.1.024

Approximations of hospital statistics of recoveries from COVID-19

A.V. Borovsky¹, A.L. Galkin², S.S. Doroshenko³

¹Baikal State University, Irkutsk, the Russian Federation

²A.M. Prokhorov Institute of General Physics of the Russian Academy of Sciences, Moscow, the Russian Federation

³Irkutsk Order of the Badge of Honor Regional Clinical Hospital, Irkutsk, the Russian Federation

Abstract. Hospital statistics on COVID-19 recoveries in Irkutsk are presented in the form of the rate of recovery over a certain number of days from the full group of patients. The recovery time varies from 1

to 182 days. The number of cases considered reaches ~ 100000 cases. For the convenience of using the data, it is proposed to approximate the table for the recovery rate by various types of nonlinear functions. The following variants of approximating functions have been studied: Gaussian, Lorentz, modified Lorentz, Weibull function, Johnson functions. For comparison with statistics, methods were used to minimize the standard deviations of approximating functions from experimental data. The least squares method is used for functions with two and three parameters, the coordinate descent method, and the gradient descent method for functions with four fitting parameters. It is shown that the best fitting results are provided by a modified Lorentz function with four parameters. According to the degree of discrepancy with experimental statistics, the approximating functions are arranged in the following order: the Weibull function provides the least accurate fit (16.15%), followed by the Johnson function S_U (10.65%), slightly better fit for the Johnson function S_B (8.49%), for the Gaussian function (5.8%), for the Lorentz function the fit is (3.2828%), the best fit is given by the modified Lorentzian function (3.2804%) under certain approximations.

Keywords: epidemic theory, optimization methods, coordinate descent, gradient descent, least squares method, Gauss approximation, Lorentz approximation, Weibull approximation, Johnson approximation, modified Lorentz distribution.

For citation: Borovsky A.V., Galkin A.L., Doroshenko S.S. Approximations of hospital statistics of recoveries from COVID-19. *Modeling, optimization, and information technology*. 2025;13(1). (In Russ.). URL: <https://moitvvt.ru/ru/journal/pdf?id=1812> DOI: 10.26102/2310-6018/2025.48.1.024

Введение

Математическое описание массовых эпидемий интересует исследователей в связи с необходимостью для государственной власти и системы здравоохранения управлять эпидемическим процессом. Некоторые достижения в области моделирования эпидемий на 2010 год изложены в работе [1]. Применительно к заболеванию COVID-19 на важные особенности эпидемического процесса, связанные с протяженным инкубационным периодом заболевания, и распространением инфекции с пассажирскими потоками, было указано в работе [2]. Дискретная модель развития эпидемии COVID-19 для большого региона представлена в работе [3]. Попытка моделирования эпидемии при помощи агентно-ориентированного подхода описана в работе [4]. Моделирование эпидемии COVID-19 при помощи уравнений эпидемической кинетики с запаздыванием при наличии квазистационарных источников заражения предпринято в [5]. Эпидемические модели должны опираться на исследования скоростей выздоровления граждан [6].

В связи с эпидемическими исследованиями [7, 8] Областная больница города Иркутска предоставила авторам обезличенные данные по госпитальному выздоровлению граждан от заболевания COVID-19 за 2020–2021 гг. Таблицы представляли собой дискретную случайную величину $t(n)$, $n = 1, 2, \dots, N$. Где n – порядковый номер пациента, поступившего на госпитальное лечение, а t – время нахождения на больничной койке от поступления до выписки из больницы в сутках. Время лечения t колебалось от одних суток до 182 суток, т. е. представляло собой дискретную случайную величину. Количество рассмотренных случаев выздоровления достигало $N = 98874$.

Дискретная случайная величина $t(n)$ обладает дискретной же плотностью, которая описывает скорость выздоровления граждан из полной группы заболевших $N = 98874$, т. е. количество граждан, которые были выписаны за сутки с номером t . Эта дискретная скорость выздоровлений на рисунках, приведенных ниже, обозначена точками. Нахождение указанной дискретной скорости выздоровлений осуществлялось из исходной таблицы $t(n)$ подсчетом числа выздоровлений, имевших одно и то же значение t .

Существенным обстоятельством является то, что количество значений t достигает 182. Поэтому использовать такую таблицу в расчетах различных средних неудобно. В

связи с этим хотелось бы найти более удобную и достаточно простую аналитическую аппроксимацию для скорости выздоровлений. Обозначим такую функцию следующим образом:

$$\frac{dn}{dt}(t) \equiv \dot{n}. \quad (1)$$

Обратим внимание, что

$$\sum_{t=1}^{182} \frac{dn}{dt}(t) = N.$$

Делая аппроксимацию дискретной скорости выздоровлений непрерывной функцией, переходим к непрерывному описанию. Это законно в связи с большим количеством госпитальных случаев $N \sim 100000$.

При непрерывном описании

$$\begin{aligned} \frac{dn}{dt}(t) &\equiv Nf(t), \quad f(t) = \frac{1}{N} \frac{dn}{dt}(t), \\ \int_0^{\infty} \frac{dn}{dt}(t) dt &= N \int_0^{\infty} f(t) dt = N, \quad \int_0^{\infty} f(t) dt = 1, \end{aligned}$$

здесь $f(t)$ имеет смысл непрерывной плотности вероятности выздоровлений.

В данной работе исследуем возможность аппроксимации дискретной скорости выздоровлений следующими функциями:

- 1) Гауссовой [7],
- 2) Лоренцевой [7],
- 3) Вейбулла [9],
- 4) Три варианта Джонсона [10],
- 5) Модифицированной Лоренцевой.

Целью данной работы является разработка улучшенных математических моделей, описывающих выздоровление от коронавирусной инфекции COVID-19 и опирающихся на экспериментальные данные.

Задачей исследования является получение наилучшей аппроксимации эмпирических данных по выздоровлениям от коронавирусной инфекции COVID-19 для госпитальных случаев по городу Иркутску за 2021–2022 гг. с помощью нескольких методов минимизации среднеквадратичных отклонений различных аппроксимирующих кривых от экспериментальных данных.

Материалы и методы

Выпишем последовательно пять указанных выше случаев аппроксимации и, соответственно, 7 видов аппроксимирующих функций. (Джонсоном предложено три вида функций.) Представим также методы оптимизации среднеквадратичного отклонения, которые применялись в данной работе.

Аппроксимация гауссовой функцией.

$$Nf_G(t; b_1, c_1, d_1) = b_1 e^{-\left(\frac{t-c_1}{d_1}\right)^2}. \quad (2)$$

Эта функция обладает тремя параметрами b_1, c_1, d_1 , которые могут быть найдены методом наименьших квадратов, минимизирующим сумму квадратичных отклонений кривой (2) от дискретной скорости выздоровлений (1). Напомним, что b_1 – значение функции (2) в максимуме, c_1 – положение максимума на оси времени t , d_1 – экспоненциальная полуширина функции (2).

Каноническая запись плотности вероятности случайной величины нормального распределения имеет следующий вид:

$$f_N(t; \bar{t}, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\bar{t}}{\sigma}\right)^2},$$

где \bar{t} – среднее значение, а σ – дисперсия случайное величины. Связь между константами аппроксимирующей функции (2) и плотности вероятности нормального распределения следующая:

$$\sigma\sqrt{2} = d_1, \quad \frac{b_1}{N} = \frac{1}{\sigma\sqrt{2\pi}}.$$

Перемножая равенства, получим, что для нормального распределения должно выполняться равенство

$$\frac{b_1 d_1}{N} \sqrt{\pi} = 1.$$

Аппроксимация Лоренцевой функцией.

$$N f_L(t; b_2, c_2, d_2) = \frac{b_2}{1 + \left(\frac{t-c_2}{d_2}\right)^2}. \quad (3)$$

Эта функция также обладает тремя параметрами b_2, c_2, d_2 , где b_2 – значение функции в максимуме, c_2 – положение максимума, d_2 – полуширина Лоренцева распределения на полувывоте.

Аппроксимация функцией Вейбулла. В теории вероятностей широко используется распределение вероятности Вейбулла [9]:

$$F(x) = 1 - \exp\left\{-\left(\frac{x-\mu}{\alpha}\right)^\beta\right\}, \quad (4)$$

где α – параметр масштаба, β – параметр формы, μ – параметр сдвига. Дифференцируя, находим функцию, описывающую плотность вероятности Вейбулла:

$$f(x) = \frac{dF(x)}{dx} = \left(\frac{\beta}{\alpha}\right) \left(\frac{x-\mu}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-\mu}{\alpha}\right)^\beta}. \quad (5)$$

Распределение и плотность вероятностей Вейбулла широко используются в метеорологии [11], химии [12], прочности материалов [13], радиофизике [14], электронике [15], медицине в радиологии [16], в моделях выздоровления от заболеваний [17], в моделях дожития, используемых в страховом деле [18], инженерии, электротехнике, гидрологии, медицинских исследованиях.

Представляет интерес аппроксимировать дискретную скорость выздоровления пациентов функцией с четырьмя константами:

$$N f_V(t; \alpha, \beta, \mu, A) = N_m A \left(\frac{\beta}{\alpha}\right) \left(\frac{t-\mu}{\alpha}\right)^{\beta-1} e^{-\left(\frac{t-\mu}{\alpha}\right)^\beta}, \quad (6)$$

где $N_m = 8728$ – максимальная скорость выздоровления, а константа $A \sim 1$ (введена, чтобы увеличить возможности аппроксимации).

Аппроксимации Джонсона. Функции Джонсона [10] обладают четырьмя параметрам $\varepsilon, \mu_1, \mu_2, \mu_3$, что придает задаче подгонки под экспериментальные точки значительную гибкость. Функции Джонсона используются в теории вероятностей [19] и математической статистике [20].

Функция Джонсона S_U . Проведем приближение экспериментальных данных функцией Джонсона S_U :

$$f_U(t) = \frac{\mu_1}{\sqrt{2\pi}} \frac{1}{\sqrt{(t-\varepsilon)^2 + (\mu_2)^2}} \exp\left\{-\frac{1}{2}\left(\mu_3 + \mu_1 \ln\left\{\frac{t-\varepsilon}{\mu_2} + \left[\left(\frac{t-\varepsilon}{\mu_2}\right)^2 + 1\right]^{1/2}\right\}\right)^2\right\}, \quad (7)$$

$$t < \infty, -\infty < \varepsilon < \infty, \mu_1 > 0, \mu_2 > 0, -\infty < \mu_3 < \infty.$$

Функция Джонсона S_B :

$$f_B(t) = \frac{\mu_1}{\sqrt{2\pi}} \frac{\mu_2}{(t-\varepsilon)(\mu_2-t+\varepsilon)} \exp \left\{ -\frac{1}{2} \left(\mu_3 + \mu_1 \ln \left\{ \frac{t-\varepsilon}{(\mu_2-t+\varepsilon)} \right\} \right)^2 \right\}, \varepsilon \leq t \leq \varepsilon + \mu_2. \quad (8)$$

Функция Джонсона S_L :

$$f_{SL}(t) = \frac{\mu_1}{\sqrt{2\pi}} \frac{1}{(t-\varepsilon)} \exp \left\{ -\frac{\mu_1^2}{2} \left(\frac{\mu_3 - \mu_1 \ln \mu_2}{\mu_1} + \ln \{t - \varepsilon\} \right)^2 \right\}, \quad (9)$$

$$\mu_1 > 0, t \geq \varepsilon, \mu_2 > 0, -\infty < \varepsilon < \mu_1, -\infty < \mu_3 - \mu_1 \ln \mu_2 < \infty, -\infty < \mu_3 < \infty.$$

Обобщенное распределение Лоренца:

$$f(t) = \frac{b}{1 + \left| \frac{t-c}{d} \right|^\gamma} \quad (10)$$

обладает четырьмя параметрами $\{b, c, d, \gamma\}$.

Критерием качества аппроксимации выберем отношение среднеквадратичного отклонения аппроксимирующей функции от дискретных экспериментальных данных к максимальному значению дискретной плотности вероятности выздоровлений:

$$\overline{\rho^2} = \frac{1}{N} \sum_{i=1}^N (f_3(t_i) - f_a(t_i))^2, \rho = \sqrt{\overline{\rho^2}}, \delta = \frac{\rho}{f_{3,max}}, \quad (11)$$

$$f_3(t_i) = \frac{1}{N} \frac{dn}{dt}(t_i), \quad f_a(t_i) = f_P(t_i; b_1, c_1, d_1),$$

где P обозначает вид аппроксимирующей функции. Если, например, используется Гауссова функция, то $P = G$.

В работе использовались несколько методов минимизации функции невязки (11). Первым был метод наименьших квадратов. Он хорошо известен, приводить его формулы здесь не будем. Этот метод применялся прежде всего для нахождения трех параметров подгонки для Гауссовой (2) и Лоренцевой (3) функций.

Второй метод назовем комбинированным методом наименьших квадратов. Он был применен для нахождения 3-х и 4-х параметров подгонки для функции Вейбулла (5) и (6). Этот метод заключался в следующем. Выбиралась сетка по параметру β . В каждой точке сетки проводилась минимизация невязки (11) методом наименьших квадратов по двум параметрам α и μ для функции (5) и трем параметрам A , α и μ для функции (6).

Попытка аппроксимации функциями Джонсона представлена авторами в работе [8] логико-эмпирическим методом исследования. В данной работе для нахождения подгоночных параметров функций Джонсона применялись строгие методы оптимизации. Использовался комбинированный подход. Сначала применялся метод покоординатного спуска [21]. Фиксировались 3 параметра, находился локальный минимум невязки на сетке значений четвертого параметра. Далее процедура повторялась для всех параметров циклически. Другая часть комбинированного подхода состояла в использовании метода спуска по градиенту [21]. Последовательно вносилось возмущение в каждый параметр и численно определялся градиент $\nabla \delta$, далее на одномерной сетке вдоль градиента определялся минимум одномерной задачи

$$\delta_{\min}^{n+1} = \min_h \delta \left(\varepsilon^{(n)} - h \frac{\partial \delta^{(n)}}{\partial \varepsilon}, \mu_1^{(n)} - h \frac{\partial \delta^{(n)}}{\partial \mu_1}, \mu_2^{(n)} - h \frac{\partial \delta^{(n)}}{\partial \mu_2}, \mu_3^{(n)} - h \frac{\partial \delta^{(n)}}{\partial \mu_3} \right) \quad (12)$$

и тем самым $(n+1)$ -ые значения параметров. Здесь минимизируется δ поскольку ρ^2 имеет большие порядки, что сказывается на точности вычисления градиента.

Результаты

В данном разделе представлены результаты решения задач минимизации для вышеуказанных аппроксимирующих функций (2), (3), (5)–(10).

Результаты для Гауссовой (2) и Лоренцевой (3) функций приведены в Таблице 1. Предварительные расчеты были выполнены в статье [7]. Эти функции обладают тремя подгоночными параметрами, которые найдены методом наименьших квадратов.

Таблица 1 – Параметры аппроксимации для Гауссовой и Лоренцевой функций
Table 1 – Approximation parameters for Gaussian and Lorentz functions

	b	c	d	δ
G	7618,61	11,4359	6,7879	0,0559
L	8652,67	11,478	4,35778	0,0328

В Таблице 2 и на Рисунке 1 представлены результаты минимизации невязки для функции Вейбулла комбинированным методом наименьших квадратов. У кривых на Рисунке 1 проставлены числовые значения β . Кривые на Рисунке 1 соответствуют следующим вариантам из Таблицы 2: 1 (черная), 2 (синяя), 3 (красная), 4 (зеленая). У кривых проставлены числовые значения β .

Таблица 2 – Параметры аппроксимации для функции Вейбулла
Table 2 – Approximation parameters for the Weibull function

№	β	A	α	μ	ρ	δ
1	2,00	0,9958	12,96	0,9166	$1,353 \times 10^4$	0,1682
2	2,40	0,9435	12,89	0,5820	$1,332 \times 10^4$	0,1655
3	2,71	0,9315	13,68	-0,3939	$1,333 \times 10^4$	0,1656
4	3,00	0,9258	14,60	-1,444	$1,339 \times 10^4$	0,1664
5	2,169	1	13,03	0,7153	$1,321 \times 10^4$	0,1641
6	2,400	1	13,53	0,0527	$1,304 \times 10^4$	0,1620
7	2,650	1	14,31	0,8778	$1,300 \times 10^4$	0,1615
8	3,840	1	19,17	-6,1423	$1,329 \times 10^4$	0,1651

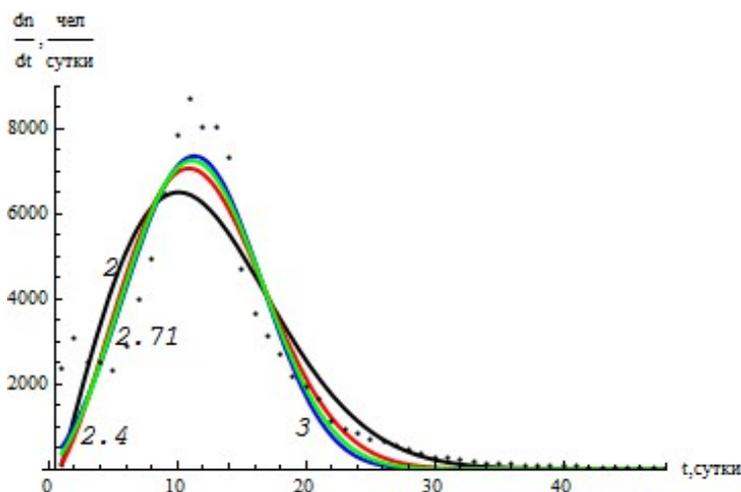


Рисунок 1 – Аппроксимация скорости выздоровления граждан с использованием функции Вейбулла (7)

Figure 1 – Approximation of the rate of recovery of citizens using the Weibull function (7)

Данные расчеты приводят к следующей зависимости эффективной погрешности δ от параметра формы (степени) β (Рисунок 2).

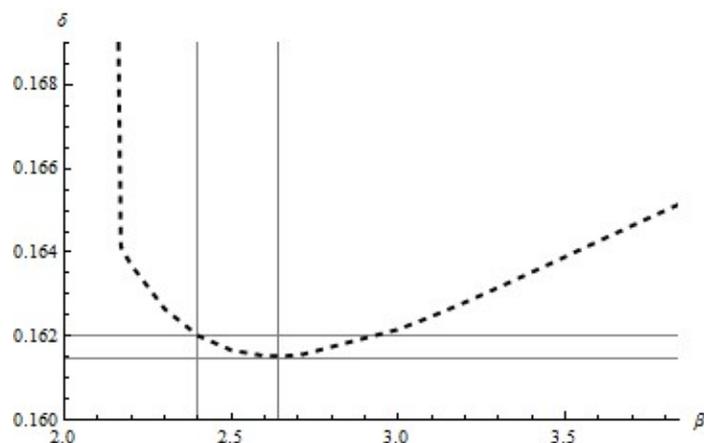


Рисунок 2 – Эффективное отклонение δ как функция параметра формы β функции Вейбулла при оптимизации по двум параметрам масштаба α и сдвига μ методом наименьших квадратов
Figure 2 – Effective deviation δ as a function of the shape parameter β of the Weibull function when optimizing for two scale α and shift μ parameters using the least squares method

Минимум невязки достигается при $\beta = 2,65$ (Рисунок 3).

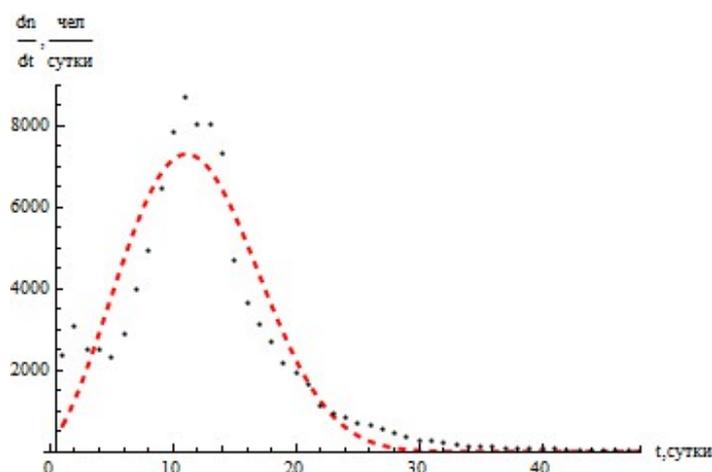


Рисунок 3 – Вариант подгонки функцией Вейбулла с минимумом невязки (Таблица 2, строка 7)
Figure 3 – A variant of the fit by the Weibull function with a minimum of deviation (Table 2, Line 7)

Приведем результаты расчетов. В Таблице 3 и на Рисунке 4 представлены результаты минимизации невязки δ для функции Джонсона S_U (7). Использовались метод покоординатного спуска, а на заключительном этапе спуск по градиенту.

Таблица 3 – Подгонка функцией Джонсона S_U (7)
Table 3 – Johnson function adjustment S_U (7)

№	μ_1	μ_2	μ_3	ε	δ
1	0,1450	0,9600	0,9000	11,50	0,12278
2	0,1675	0,8350	0,8650	10,50	0,11043
3	0,1815	1,1000	0,8810	10,52	0,10661
4	0,1805	1,1139	0,9030	10,50	0,10649
5	0,1792	1,1126	0,9017	10,50	0,10646

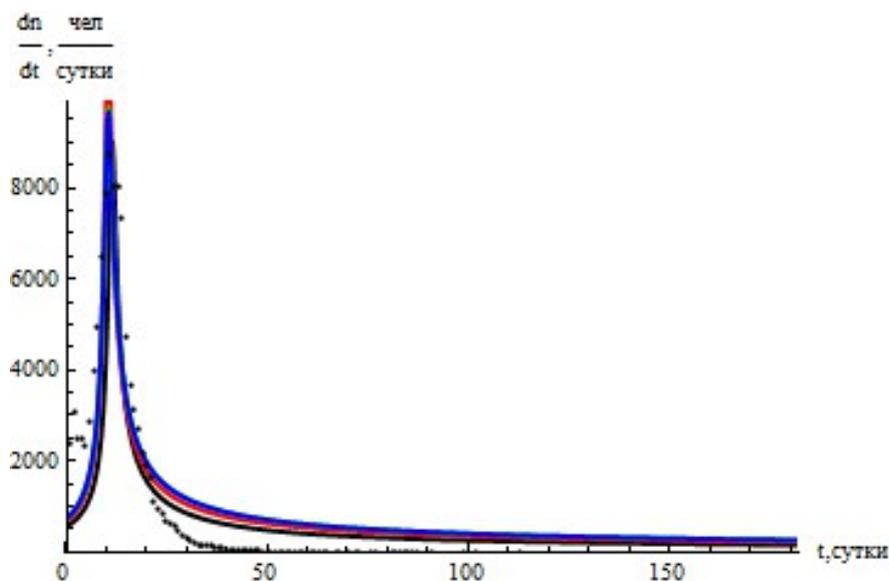


Рисунок 4 – Результаты оптимизации скорости выздоровления функцией Джонсона S_U
Figure 4 – Results of optimization of the recovery rate by the Johnson function S_U

В Таблице 4 и на Рисунке 5 представлены результаты минимизации невязки δ для функции Джонсона S_B (8). Использовались метод покоординатного спуска, а на заключительном этапе спуск по градиенту.

Таблица 4 – Подгонка функцией Джонсона S_B (8)
Table 4 – Johnson function adjustment S_B (8)

№	μ_1	μ_2	μ_3	ε	δ
1	1,6048	185	4,94349	0,0	0,19744
2	1,8663	185,0008	4,85196	0,01034	0,08527
3	1,8669	185,0007	4,85779	0,01095	0,08489

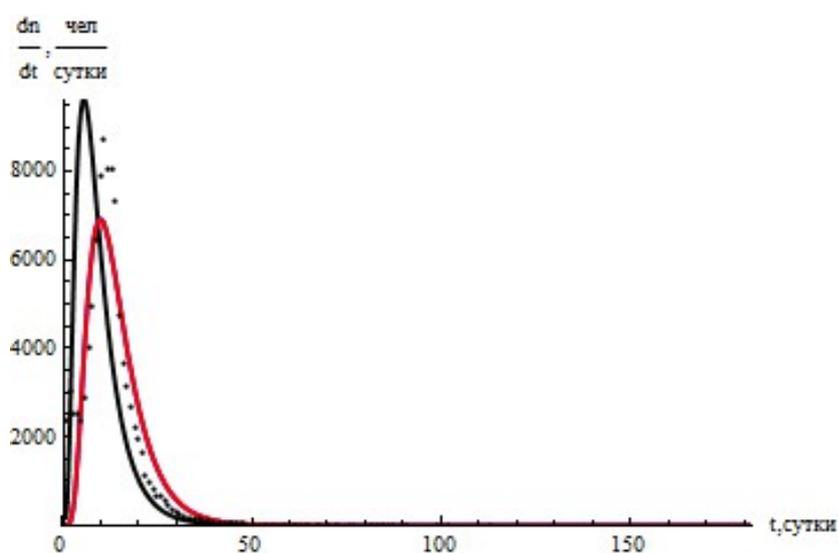


Рисунок 5 – Результаты оптимизации подгонки скорости выздоровления функцией Джонсона S_B
Figure 5 – Results of optimizing the adjustment of the recovery rate by the Johnson function S_B

На Рисунке 6 представлена попытка аппроксимации экспериментальных данных (точки) функцией Джонсона S_L .

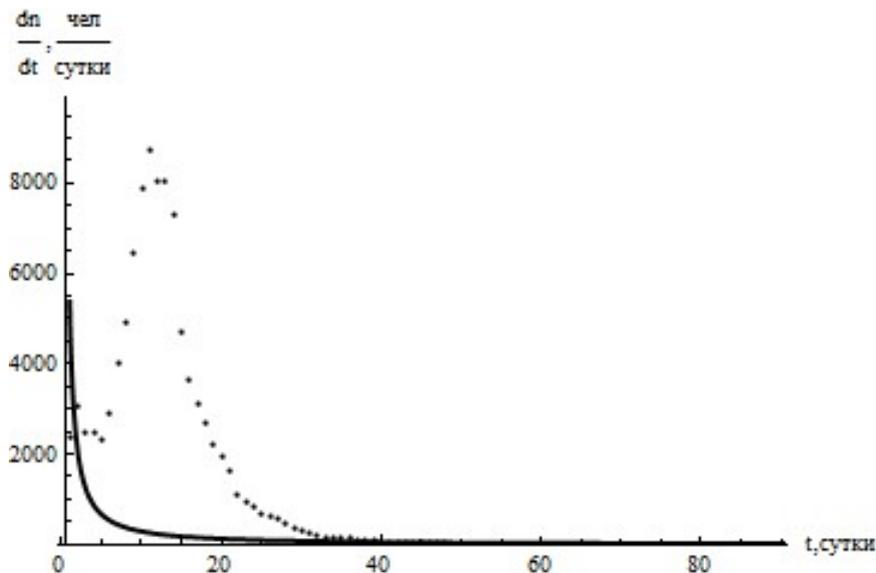


Рисунок 6 – Представлен результат трех шагов градиентного спуска для функции S_L (черная кривая)

Figure 6 – The result of three steps of gradient descent for the function S_L (black curve) is presented

В Таблице 5 и на Рисунках 7, 8 представлены результаты минимизации невязки δ для модифицированной функции Лоренца (10). Использовались метод покоординатного спуска, а на заключительном этапе спуск по градиенту.

Таблица 5 – Подгонка скоростей выздоровления обобщенным распределением Лоренца
Table 5 – Adjustment of recovery rates by generalized Lorentz distribution

№	b	c	d	γ	δ
1	7726,68	12,0259	6,02823	2,0	0,060956
2	7726,67	11,5127	5,36397	2,57440	0,042750
3	7726,67	11,4369	5,07226	2,170719	0,038315
4	7726,67	11,4527	5,00840	2,209783	0,038160
5	8652,67	11,4751	4,35751	1,978901	0,032804

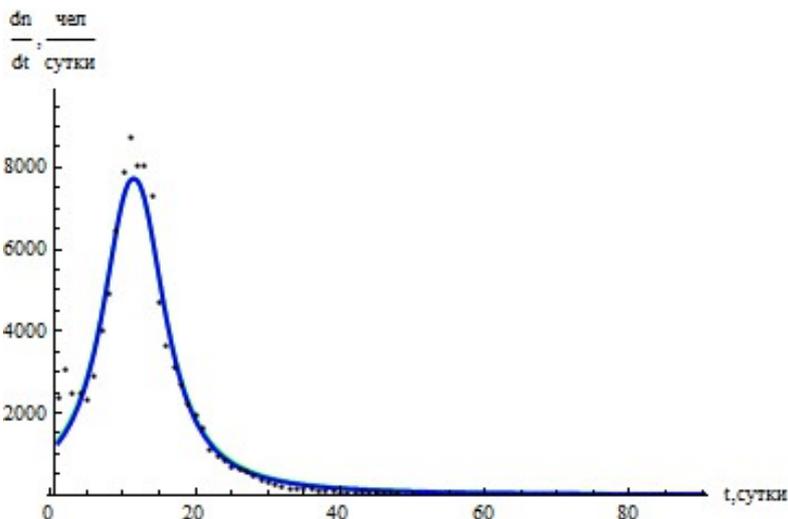


Рисунок 7 – Подгонка с использованием модифицированного распределения Лоренца (10)
Figure 7 – Fitting using a modified Lorentz distribution (10)

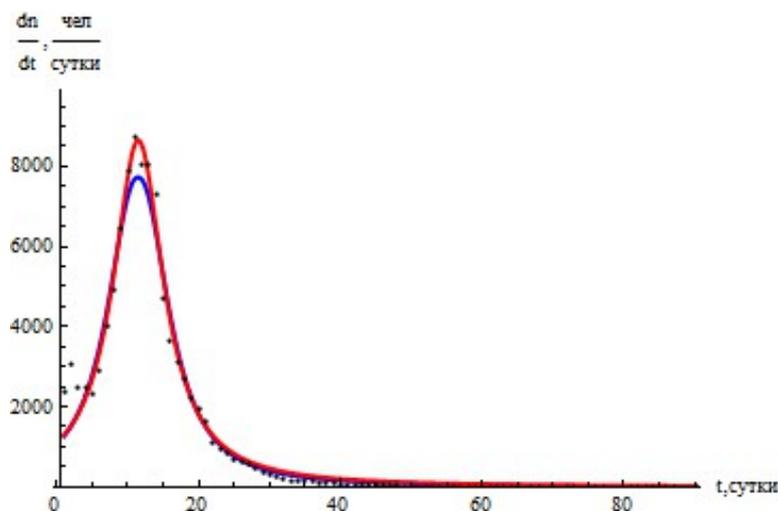


Рисунок 8 – Подгонка с использованием модифицированного распределения Лоренца (10) и разными начальными приближениями: синяя кривая (Таблица 5, строка 4), красная кривая (Таблица 5, строка 5)

Figure 8 – Fitting using a modified Lorentz distribution (10) and different initial approximations: blue curve (Table 5, row 4), red curve (Table 5, row 5)

Обсуждение

Результаты аппроксимации Гауссовой функцией, Таблица 1, показывают, что плотность вероятности выздоровлений в медицинской статистике не распределена по нормальному закону. Если взять значения для b_1 , d_1 из Таблицы 1, то вычисление дает

$$\frac{b_1 d_1}{N} \sqrt{\pi} = 0,927.$$

Отклонение от единицы составляет 7,3 %. Для нормального распределения должна быть единица. В работе [8] на основе критерия Пирсона показано, что гипотеза описания экспериментальных данных, представленных в статье, плотностью вероятности нормального распределения не состоятельна.

Оценка среднего значения из дискретных данных и аппроксимации также даёт расхождения $\bar{t} = 12,3835$, $c_1 = 11,4359$. Величина расхождения составляет 7,6 %.

Первая строка в Таблице 3 соответствует начальным данным и черной кривой на Рисунке 4. Вторая строка получена в результате многократного применения спуска по координатам (красная кривая). Третья (синяя кривая) и четвертая (зеленая кривая сливается с синей) строки получены в результате градиентного спуска. Малое отличие третьей и четвертой строк свидетельствует о завершении минимизации. Как следует из рисунков, при минимизации невязки идет подстройка в районе максимума и средней части распределения, но на правом крыле происходит даже удаление от экспериментальных данных. Аппроксимация лучше, чем в варианте распределения Вейбулла, но хуже Лоренцевой функции.

Аппроксимация медицинской статистики функцией Вейбулла (Таблица 2, Рисунок 1) выполнена комбинированным методом наименьших квадратов.

Аппроксимация статистических данных функциями Джонсона S_U, S_B проводилась комбинацией методов покоординатного спуска и на заключительном этапе спуска по градиенту. Результаты расчетов представлены в Таблицах 3, 4 и на Рисунках 4, 5. Применение функции Джонсона S_L не дает адекватных результатов, на что указывает Рисунок 6.

Первые 5 чисел в Таблице 4 соответствуют начальным данным и черной кривой на Рисунке 5. Вторые 5 чисел (красная кривая) и третьи 5 чисел (синяя кривая, совпадает с красной) получены в результате двух шагов градиентного спуска.

Малое отличие второй и третьей пятерок чисел свидетельствует о завершении минимизации. Метод минимизации невязки приводит к результатам, аналогичным результатам применения метода на основе статистических оценок. При этом величина ε не сильно отличается от 0, так что нет необходимости проведения дополнительных исследований. Аппроксимация лучше, чем в варианте распределения Вейбулла, но хуже Лоренцевой и Гауссовой функций.

С позиции нормированных показателей асимметрии и эксцесса приемлемость семейств Джонсона для аппроксимации скоростей выздоровления в медицине располагается в последовательности S_U, S_B, S_L . С позиции достижимости минимума невязки, приемлемость последовательности кривых Джонсона следующая S_B, S_U, S_L .

Представляет интерес аппроксимация медицинской статистики модифицированной функцией Лоренца (10). Начальное приближение выбирается достаточно произвольно (первая строка в Таблице 5), вторая строка – один шаг градиентного спуска для модифицированной функции Лоренца. Еще один шаг градиентного спуска – третья строка в Таблице 5. Эффективная погрешность 3,83 %, заметна подстройка в средней части за счет изменения d, γ . Еще один шаг градиентного спуска – четвертая строка в Таблице 5. Синяя кривая, Рисунок 8, эффективная погрешность 3,82 %. Старт с произвольного начального приближения приводит к малым значениям эффективной погрешности, но эти значения больше погрешности аппроксимации Лоренцевой функцией. Если в качестве начального приближения использовать параметры обычного распределения Лоренца из второй строки Таблицы 1, то обобщенное распределение Лоренца за один шаг градиентного спуска (пятая строка в Таблице 5) позволяет еще улучшить аппроксимацию на больничной статистике по сравнению с обычным распределением Лоренца.

Заключение

В работе рассмотрена аппроксимация экспериментальной скорости выздоровления от COVID-19 различными видами аппроксимирующих функций. Целью является найти наиболее подходящий вид аппроксимации с наименьшим значением среднеквадратичного отклонения между экспериментальными точками и непрерывной функцией. Рассмотрены следующие виды аппроксимирующих функций: Гауссова, с тремя подгоночными параметрами (2); Лоренцева, с тремя подгоночными параметрами (3); функция Вейбулла, с тремя и четырьмя подгоночными параметрами (5), (6); функции Джонсона S_U, S_B, S_L с четырьмя подгоночными параметрами (7), (8) и (9); модифицированная функция Лоренца с четырьмя подгоночными параметрами (10). В качестве методов оптимизации использовались: метод наименьших квадратов для функций с двумя и тремя подгоночными параметрами, комбинированный метод наименьших квадратов для функций с тремя и четырьмя параметрами, метод спуска по координатам и метод спуска по градиенту для функций с четырьмя параметрами.

Исследование показало, что функции, перечисленные выше, располагаются в следующей последовательности: функция Вейбулла обеспечивает наименее точную подгонку (16,15 %), далее располагается функция Джонсона S_U (10,65 %), несколько лучше подгонка для функции Джонсона S_B (8,49 %), для Гауссовой функции (5,8 %), для Лоренцевой функции подгонка составляет (3,2828 %), наилучшую подгонку дает модифицированная Лоренцева функция (3,2804 %). Эта функция описывается формулой (10). Параметры наилучшей подгонки даны в строке 5 Таблицы 5.

Аппроксимирующая функция может быть использована для решения многих медико-математических задач. Например, для вычисления среднего времени лечения в стационарах, для определения средних затрат на лечение больных, для нахождения ядра интегрального оператора в модели эпидемической кинетики для города [22].

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Лопатин А.А., Сафронов В.А., Раздорский А.С., Куклев Е.В. Современное состояние проблемы математического моделирования и прогнозирования эпидемического процесса. *Проблемы особо опасных инфекций*. 2010;(3):28–30. Lopatin A.A., Safronov V.A., Razdorskiy A.S., Kouklev E.V. Current State of Problem of Mathematical Modeling and Prognosis of the Epidemic Process. *Problems of Particularly Dangerous Infections*. 2010;(3):28–30. (In Russ.).
2. Головинский П.А. Математическое моделирование распространения вирусов с длинной инкубационной фазой в тесном мире. *Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии*. 2020;(2):5–14. <https://doi.org/10.17308/sait.2020.2/2909> Golovinski P.A. Mathematical modelling of the transmission of viruses with a long incubation period in a small-world network. *Proceedings of Voronezh State University. Series: Systems Analysis and Information Technologies*. 2020;(2):5–14. (In Russ.). <https://doi.org/10.17308/sait.2020.2/2909>
3. Trigger S.A., Czerniawski E.B. Equation for Epidemic Spread with the Quarantine Measures: Application to COVID-19. *Physica Scripta*. 2020;95(10). <https://doi.org/10.1088/1402-4896/abb2e2>
4. Макаров В.Л., Бахтизин А.Р., Сушко Е.Д., Агеева А.Ф. Моделирование эпидемии COVID-19 – преимущества агент-ориентированного подхода. *Экономические и социальные перемены: факты, тенденции, прогноз*. 2020;13(4):58–70. <https://doi.org/10.15838/esc.2020.4.70.3> Makarov V.L., Bakhtizin A.R., Sushko E.D., Ageeva A.F. COVID-19 Epidemic Modeling – Advantages of an Agent-Based Approach. *Ekonomicheskie i Sotsial'nye Peremeny: Fakty, Tendentsii, Prognoz*. 2020;13(4):58–70. (In Russ.). <https://doi.org/10.15838/esc.2020.4.70.3>
5. Borovsky A.V., Galkin A.L. Model of Epidemic Kinetics with a Source on the Example of Moscow. *Computational and Mathematical Methods in Medicine*. 2022;2022. <https://doi.org/10.1155/2022/6145242>
6. Цветков В.В., Токин И.И., Лиознов Д.А., Венев Е.В., Куликов А.Н. Прогнозирование длительности стационарного лечения пациентов с COVID-19. *Медицинский совет*. 2020;(17):82–90. <https://doi.org/10.21518/2079-701X-2020-17-82-90> Tsvetkov V.V., Tokin I.I., Lioznov D.A., Venev E.V., Kulikov A.N. Predicting the duration of inpatient treatment for COVID-19 patients. *Medical Council*. 2020;(17):82–90. (In Russ.). <https://doi.org/10.21518/2079-701X-2020-17-82-90>
7. Боровский А.В., Галкин А.Л., Ильиных Н.Н., Козлова С.С. Новые результаты эпидемических моделей на примере COVID-19. *System Analysis & Mathematical Modeling*. 2022;4(4):255–274. [https://doi.org/10.17150/2713-1734.2022.4\(4\).255-274](https://doi.org/10.17150/2713-1734.2022.4(4).255-274) Borovsky A.V., Galkin A.L., Ilinykh N.N., Kozlova S.S. New Results of Epidemic Models on the Example of COVID-19. *System Analysis & Mathematical Modeling*. 2022;4(4):255–274. (In Russ.). [https://doi.org/10.17150/2713-1734.2022.4\(4\).255-274](https://doi.org/10.17150/2713-1734.2022.4(4).255-274)
8. Боровский А.В., Галкин А.Л., Козлова С.С. Аппроксимация статистических данных заболеваемости коронавирусной инфекцией с учетом расслоения по

- сопутствующим диагнозам. *Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика*. 2024;(3):95–106. <https://doi.org/10.24143/2072-9502-2024-3-95-106>
- Borovsky A.V., Galkin A.L., Kozlova S.S. Mathematical modeling of statistical data on the incidence of new coronavirus infection, taking into account the stratification by concomitant diagnoses. *Vestnik of Astrakhan State Technical University. Series: Management, Computer Science and Informatics*. 2024;(3):95–106. (In Russ.). <https://doi.org/10.24143/2072-9502-2024-3-95-106>
9. Королев В.Ю., Соколов И.А. Об условиях сходимости распределений экстремальных порядковых статистик к распределению Вейбулла. *Информатика и ее применения*. 2014;8(3):3–11. <https://doi.org/10.14357/19922264140301>
- Korolev V.Yu., Sokolov I.A. On conditions of convergence of the distributions of extremal order statistics to the Weibull distribution. *Informatics and Applications*. 2014;8(3):3–11. (In Russ.). <https://doi.org/10.14357/19922264140301>
10. Johnson N.L. Systems of Frequency Curves Generated by Methods of Translation. *Biometrika*. 1949;36(1/2):149–176. <https://doi.org/10.2307/2332539>
11. Иванова Ю.П., Соколова Е.В., Сахарова А.А., Иванова О.О., Азаров В.Н. Проверка выполнения закона Вейбулла для различных направлений ветра, характерных для линейного города Волгограда. *Вестник Волгоградского государственного архитектурно-строительного университета. Серия: Строительство и архитектура*. 2020;(3):134–141.
- Ivanova Yu.P., Sokolova E.V., Sakharova A.A., Ivanova O.O., Azarov V.N. Checking compliance with Weibull's law for various wind directions typical of the linear city of Volgograd. *Vestnik Volgogradskogo gosudarstvennogo arhitekturno-stroitel'nogo universiteta. Seriya: Stroitel'stvo i arhitektura*. 2020;(3):134–141. (In Russ.).
12. Бойко Ю.М., Марихин В.А., Мясникова Л.П., Москалюк О.А., Радованова Е.И. Статистический анализ прочности ультраориентированных пленочных нитей сверхвысокомолекулярного полиэтилена в рамках модели Вейбулла. *Физика твердого тела*. 2016;58(10):2065–2068.
- Boiko Yu.M., Marikhin V.A., Myasnikova L.P., Moskalyuk O.A., Radovanova E.I. Statistical analysis of the strength of ultra-oriented ultra-high-molecular-weight polyethylene film filaments in the framework of the Weibull model. *Physics of the Solid State*. 2016;58(10):2141–2144. <https://doi.org/10.1134/S1063783416100103>
13. Прохоров С.А., Даниленко М.С. Модель прогнозирования дефектных участков магистральных газопроводов с помощью заданного закона распределения Вейбулла. *Естественные и технические науки*. 2016;(4):220–224.
14. Гродзенская И.С. Исследование эффективности последовательных методов обнаружения сигналов на фоне помех, имеющих распределение Вейбулла. *Метрология*. 2006;(7):30–35.
15. Шнейдеров Е.Н. Использование распределения Вейбулла для группового прогнозирования параметрической надёжности изделий электронной техники. В сборнике: *Современные средства связи: Материалы XVII Международной научно-технической конференции, 16–18 октября 2012 года, Минск, Беларусь*. Минск: Высший государственный колледж связи; 2012. С. 152–153.
16. Осовец С.В., Азизова Т.В., Гергенрейдер С.Н. Методы оценки неопределенности дозовых порогов для детерминированных эффектов. *Медицинская радиология и радиационная безопасность*. 2010;55(3):11–16.
- Osovetz S.V., Azizova T.V., Gergenreider S.N. Methods of Uncertainty Assessment for Deterministic Effects Dose Thresholds. *Medical Radiology and Radiation Safety*. 2010;55(3):11–16. (In Russ.).

17. Elandt-Johnson R.C., Johnson N.L. *Survival Models and Data Analysis*. New York: John Wiley & Sons, Inc.; 1999. 457 p.
18. Hogg R.V., Klugman S.A. *Loss distributions*. New York: John Wiley & Sons, Inc.; 1984. 235 p.
19. Кобзарь А.И. *Прикладная математическая статистика. Для инженеров и научных работников*. Москва: ФИЗМАТЛИТ; 2006. 816 с.
20. Борбач Н.М., Школина Т.В. Процедура подбора кривой из системы Джонсона методами процентилей и максимального правдоподобия – наименьших квадратов в R. *System Analysis & Mathematical Modeling*. 2023;5(4):476–493. [https://doi.org/10.17150/2713-1734.2023.5\(4\).476-493](https://doi.org/10.17150/2713-1734.2023.5(4).476-493)
Borbats N.M., Shkolina T.V. The Procedure for Selecting a Curve from the Johnson System by Percentile Matching and Maximum Likelihood and Least Squares Approaches in R. *System Analysis & Mathematical Modeling*. 2023;5(4):476–493. (In Russ.). [https://doi.org/10.17150/2713-1734.2023.5\(4\).476-493](https://doi.org/10.17150/2713-1734.2023.5(4).476-493)
21. Калиткин Н.Н. *Численные методы*. Москва: Главная редакция физико-математической литературы изд-ва «Наука»; 1978. 512 с.
22. Боровский А.В., Дорошенко С.С. Вывод выражения для ядра интегрального оператора в интегро-дифференциальной модели распространения эпидемии COVID-19. В сборнике: *Ляпуновские чтения – 2024: Материалы 40-й международной конференции, 02–06 декабря 2024 года, Иркутск, Россия*. Иркутск: ИДСТУ СО РАН; 2024. С. 26–29.

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Боровский Андрей Викторович, доктор физико-математических наук, профессор кафедры математических методов и цифровых технологий, Байкальский государственный университет, Иркутск, Российская Федерация.

e-mail: andrei-borovskii@mail.ru

ORCID: [0000-0003-2119-1072](https://orcid.org/0000-0003-2119-1072)

Andrey V. Borovsky, Doctor of Physical and Mathematical Sciences, Professor, Department of Mathematical Methods and Digital Technologies, Baikal State University, Irkutsk, the Russian Federation.

Галкин Андрей Леонидович, доктор физико-математических наук, старший научный сотрудник, Институт общей физики им. А.М. Прохорова РАН, Москва, Российская Федерация.

e-mail: galkin@kapella.gpi.ru

ORCID: [0000-0002-1132-9202](https://orcid.org/0000-0002-1132-9202)

Andrey L. Galkin, Doctor of Physical and Mathematical Sciences, Senior Researcher, A.M. Prokhorov Institute of General Physics of the Russian Academy of Sciences, Moscow, the Russian Federation.

Дорошенко Светлана Сергеевна, начальник отдела АСУ, Иркутская ордена «Знак почета» областная клиническая больница, Иркутск, Российская Федерация.

e-mail: kozlova_ss@iokb.ru

Svetlana S. Doroshenko, Head of the Automated Control System Department, Irkutsk Order of the Badge of Honor Regional Clinical Hospital, Irkutsk, the Russian Federation.

Статья поступила в редакцию 04.02.2025; одобрена после рецензирования 19.02.2025; принята к публикации 26.02.2025.

The article was submitted 04.02.2025; approved after reviewing 19.02.2025; accepted for publication 26.02.2025.