

## ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ОЦЕНКИ И АНАЛИЗА ПОКАЗАТЕЛЕЙ ВЫЖИВАЕМОСТИ И СМЕРТНОСТИ НАСЕЛЕНИЯ РОССИИ

Е. И. Прокудина<sup>1</sup>, А. Р. Вахитов<sup>2</sup>

<sup>1</sup>preliv@gmail.com, <sup>2</sup>ejwhfk@gmail.com

<sup>1</sup> ФГБОУ ВО «Уфимский государственный авиационный технический университет» (УГАТУ)

<sup>2</sup> ЗАО ЦСИ «Интегро»

*Поступила в редакцию 13 декабря 2021 г*

**Аннотация.** Рассматривается задача оценки и анализа числовых и функциональных характеристик продолжительности жизни новорожденного и остаточной продолжительности жизни для населения России, дифференцированного по различным признакам: возрасту, полу, региону, типу поселения, причине смерти. На основе алгоритмов вычисления характеристик продолжительности жизни по коэффициентам смертности из Российской базы данных по рождаемости и смертности Центра демографических исследований Российской экономической школы разработано программное обеспечение. В программном обеспечении реализовано построение таблиц смертности, визуализация функциональных характеристик продолжительности жизни, проведение их сравнительного анализа.

**Ключевые слова:** алгоритмы оценки показателей выживаемости и смертности; программное обеспечение; демографическая база данных; население России.

### ВВЕДЕНИЕ

Повышение ожидаемой продолжительности жизни населения является национальной целью России на 2030 год [1].

Изучение показателей выживаемости и смертности населения России, дифференцированного по различным признакам, а именно: по возрасту, полу, региону, типу поселения, причине смерти, необходимо для получения развернутой картины демографической ситуации как в России в целом, так и в отдельных ее регионах.

Задача анализа показателей, характеризующих продолжительность жизни человека, и факторов, наиболее влияющих на эти показатели, рассматривается в разных областях: в медицине, демографии, страховании, социологии. Для решения этой задачи применяются разнообразные математические модели и методы [2–7].

Демографические базы данных и информационно-аналитические системы также способствуют таким исследованиям [8, 9]. Достаточно обширную информацию о подобных ресурсах можно найти, в частности, в приложении к электронному журналу Демоскоп Weekly Института демографии НИУ ВШЭ имени А. Г. Вишневого [10].

Тем не менее, на наш взгляд, целесообразно разработать программное обеспечение, которое предоставит пользователю как подробную информацию о числовых и функциональных характеристиках продолжительности жизни для населения России, сгруппированного по различным признакам, так и возможность ее анализа. Такое программное обеспечение должно использовать базу данных. Для наших целей выбрана Российская база данных по рождаемости и смертности Центра демографических исследований Российской экономической

школы [11]. В базу включены, в частности, коэффициенты смертности по России и ее регионам, начиная с 1959 года, которые могут быть рассчитаны на основе данных официальной статистики, и которые не представлены на сайте Росстата и в международных базах данных. Основой для расчета всех показателей являются данные Российской государственной статистики [11].

Уточним, что мы понимаем под функциональными и числовыми характеристиками продолжительности жизни человека, для оценки и анализа которых разрабатывается программное обеспечение. Полагаем, что продолжительность жизни человека в возрасте  $x$  лет  $T(x)$  является непрерывной случайной величиной.  $T(0)$  – продолжительность жизни новорожденного [12].

Функциональными характеристиками случайной величины  $T(x)$  являются:  $F_x(t)$  – функция распределения;  $s_x(t)$  – функция выживания, вероятность того, что человек в возрасте  $x$  проживет еще  $t$  лет;  $f_x(t)$  – кривая смертей, плотность распределения случайной величины;  $\mu_x(t)$  – интенсивность смертности или функция риска, числовыми: математическое ожидание времени жизни, вероятности умереть в определенном интервале возрастов и т.д.

### ПОСТАНОВКА ЗАДАЧИ

Дано: Российская база данных по рождаемости и смертности Центра демографических исследований Российской экономической школы.

Требуется разработать программное обеспечение оценки и анализа показателей выживаемости и смертности населения России, дифференцированного по различным признакам, реализующее следующие функции:

- построение таблиц смертности;
- вычисление и визуализацию функциональных характеристик продолжительности жизни.

### ЗАДАЧА ПОСТРОЕНИЯ ТАБЛИЦ СМЕРТНОСТИ

Постановка задачи.

Дано:

- таблица с коэффициентами смертности населения  $m_x$  ( $x$  принимает целые неотри-

цательные значения) по регионам, годам, категориям поселения, полу, возрастным группам и причинам смерти (при наличии);

- таблица со среднегодовым населением по областям России  $N_x$  для расчета показателей смертности по  $k$ -летним группам;
- $k$ -летние возрастные группы ( $k$  равно 1 или 5).

Требуется построить таблицу смертности.

Приведем алгоритм построения полной ( $k = 1$ ) таблицы смертности. Формулы для вычисления показателей взяты в [12].

1. Выбрать параметры вычислений: территорию, год, категорию поселения, пол, причину смерти (при наличии).

2. Вычислить среднее число умерших в возрасте  $x$   $d_x$  по следующей формуле:

$$d_x = m_x \cdot N_x.$$

3. Вычислить среднее число живых в возрасте  $x$   $l_x$  по формуле:

$$l_x = l_{x+1} + d_x.$$

Вычисления следует начинать с предельного возраста  $\omega$ , для которого  $l_\omega = 0$ .

4. Нормировать значения числа доживших  $l_x$  так, чтобы значение  $l_0$  стало равно 1.

5. Умножить нормированные значения числа доживших  $l_x$  на  $10^5$ .

6. Вычислить вероятность того, что человек в возрасте  $x$  проживет еще год, обозначаемую  $p_x$ , по формуле:

$$p_x = \frac{l_{x+1}}{l_x}.$$

7. Вычислить вероятность, что человек в возрасте  $x$  скончается в течение года, обозначаемую  $q_x$ , по формуле:

$$q_x = \frac{l_x - l_{x+1}}{l_x}.$$

8. Вычислить суммарное среднее число человеко-лет  $L_x$ , прожитое членами совокупности в возрасте от  $x$  до  $x + 1$ , при предположении, что в течение года время смерти равномерно распределено, по формуле:

$$L_x = \frac{l_x + l_{x+1}}{2}.$$

9. Вычислить среднюю продолжительность предстоящей жизни  $e_x$  по формуле:

$$e_x = \frac{1}{l_x} \cdot \sum_{k=x}^{\omega} L_k.$$

10. Объединить вычисленные показатели  $l_x$ ,  $d_x$ ,  $p_x$ ,  $q_x$ ,  $L_x$ ,  $e_x$  в таблицу.

Краткая таблица смертности при  $k = 5$  строится по аналогичному алгоритму.

### ЗАДАЧА ВЫЧИСЛЕНИЯ ФУНКЦИОНАЛЬНЫХ ХАРАКТЕРИСТИК ПРОДОЛЖИТЕЛЬНОСТИ ЖИЗНИ

Постановка задачи.

Дано: таблица смертности.

Требуется найти значения и построить графики следующих функциональных характеристик продолжительности жизни новорожденного и человека, дожившего до возраста  $x$ : функций распределения, выживания, кривой смертей, интенсивности смертности.

Для решения поставленной задачи используются формулы вычисления статистических оценок функциональных характеристик продолжительности жизни [12]. Данные формулы представлены в табл. 1.

Таблица 1

#### Статистические оценки функциональных характеристик продолжительности жизни

| Функция   | Статистическая оценка     |
|---|---------------------------|
| Функция распределения продолжительности жизни новорожденного, $F(x)$                        | $1 - \frac{l_{x+1}}{l_0}$ |
| Функция выживания для новорожденного, $s(x)$  | $\frac{l_{x+1}}{l_0}$     |
| Кривая смертей для новорожденного, $f(x)$   | $\frac{d_x}{l_0}$         |
| Интенсивность смертности (функция риска) для новорожденного, $\mu(x)$                       | $\frac{d_x}{l_x}$         |
| Функция распределения остаточной продолжительности жизни человека в возрасте $x$ , $F_x(t)$ | $1 - \frac{l_{x+t}}{l_x}$ |
| Функция выживания для человека в возрасте $x$ , $s_x(t)$                                    | $\frac{l_{x+t}}{l_x}$     |
| Кривая смертей для человека в возрасте $x$ , $c$  | $\frac{d_{x+t}}{l_x}$     |
| Интенсивность смертности (функция риска) для человека в возрасте $x$ , $\mu_x(t)$           | $\frac{d_{x+t}}{l_{x+t}}$ |

Поскольку входные данные – коэффициенты смертности по причинам смерти имеются только для пятилетних групп, то есть с интервалом в 5 лет, то в программном обеспечении реализована интерполяция таких данных. Используются 3 типа интерполяционных функций, аналогичных функциям из предположений для дробных возрастов [12].

### ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

Программное обеспечение разработано на языке программирования C# (среда разработки Visual Studio 2019).

Входные данные из Российской базы данных по рождаемости и смертности Центра демографических исследований Российской экономической школы [11], представленные в формате \*.txt, включают в себя:

- перечень территорий России с кодами системы обозначения объектов административно-территориального деления (СОАТО) (для разных лет);

- краткая советская номенклатура причин смерти, основанная на ICD9;

- краткая российская номенклатура причин смерти, основанная на ICD10;

- коэффициенты смертности по возрасту (однолетние группы), категории населения и полу в России и на ее территориях за 1989–2019 гг.;

- коэффициенты смертности по причинам смерти, возрасту (пятилетние группы), категории населения и полу в России и на ее территориях за 2015–2019 гг.;

- среднегодовое население по территориям России для расчета коэффициентов смертности за 1989–2019 гг. по однолетним возрастным группам;

- среднегодовое население по территориям России для расчета коэффициентов смертности за 2015–2019 гг. по пятилетним возрастным группам.

На панели управления в левой части окна программы можно выбрать параметры вычислений: тип возрастных групп (однолетние или пятилетние), год, название региона, категория поселения (городское, сельское, все население), пол, причина смерти. Результаты вычислений отображаются в правой части в соответствующих вкладках.

Далее на рисунках представлены примеры, демонстрирующие работу программы.

На рис. 1 представлены результаты вычислений, связанных с построением краткой таблицы смертности при следующих входных параметрах: 2019 год, Республика Башкортостан, городское население, мужчины, все причины смерти. Кроме таблицы смертности выводятся графики зависимостей  $p_x$  и  $e_x$  от возраста  $x$ .

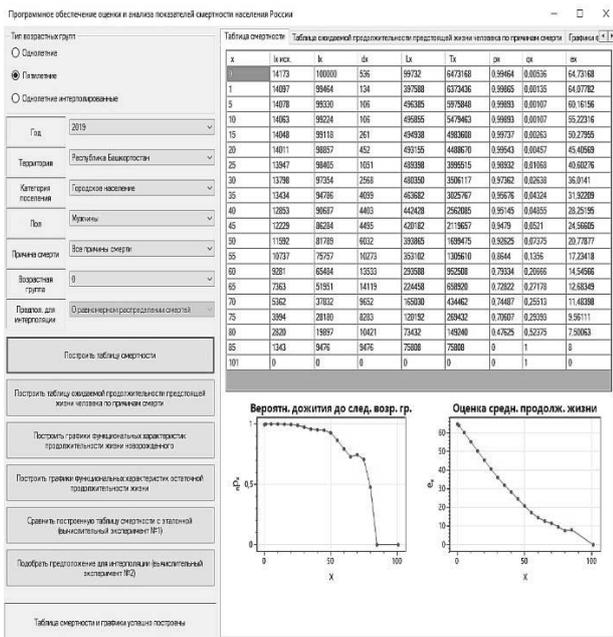


Рис. 1. Пример построения краткой таблицы смертности

На рис. 2 представлены результаты вычислений ожидаемой продолжительности предстоящей жизни человека по причинам смерти. Вычисления проведены для пятилетних возрастных групп. Входные параметры: 2019 год, Российская Федерация после 17 марта 2014 года, все население, женщины, возрастная группа – 0 (новорожденные). Значения в построенной таблице отсортированы по возрастанию продолжительности жизни. Наименьшая продолжительность жизни соответствует в данном случае причине смерти: «гипертоническая

болезнь с преимущественным поражением сердца», наибольшая – «старость».

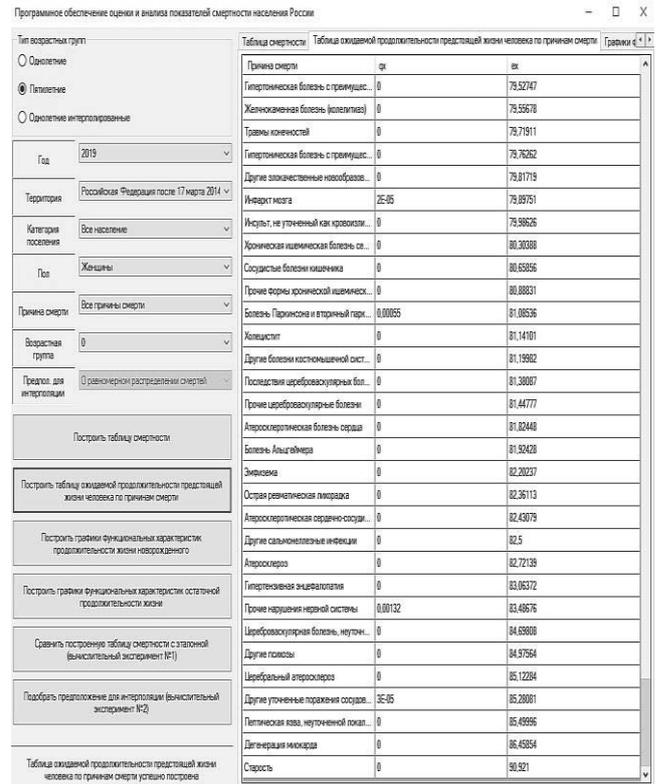
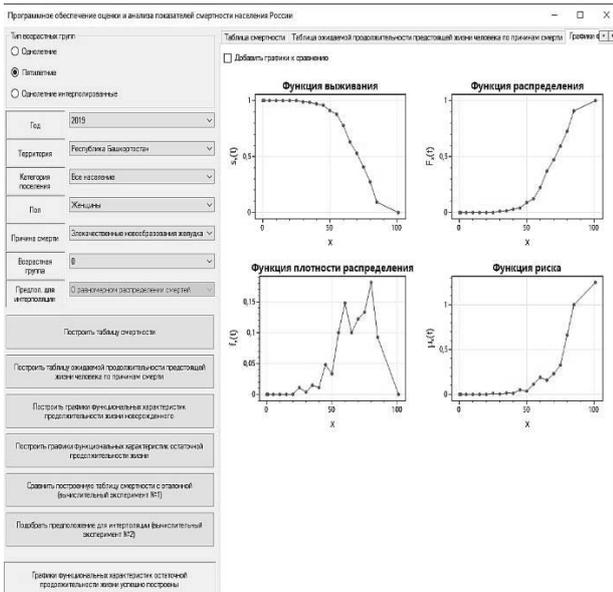


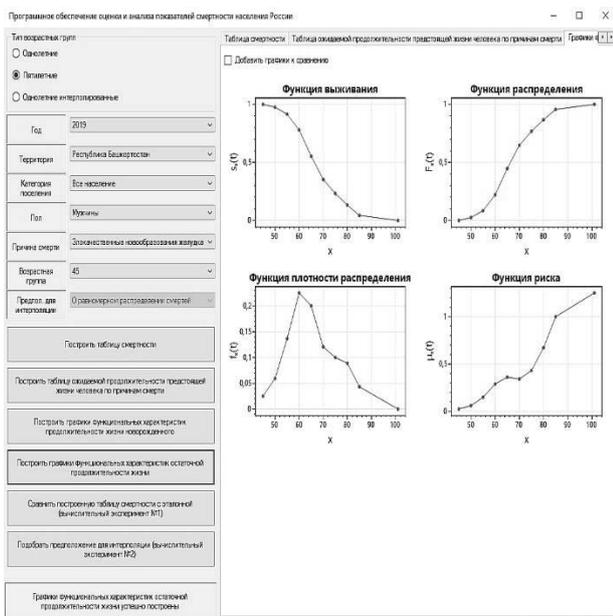
Рис. 2. Пример вычисления оценок ожидаемой продолжительности предстоящей жизни человека по причинам смерти

На рис. 3 представлен пример построения графиков статистических оценок функциональных характеристик продолжительности жизни новорожденного. Для вычислений выбраны пятилетние группы возраста. Входные параметры: 2019 год, Республика Башкортостан, все население, женщины, причина смерти – злокачественные новообразования желудка.

На рис. 4 представлен пример построения графиков статистических оценок функциональных характеристик остаточной продолжительности жизни. Для вычислений выбраны пятилетние группы возраста. Входные параметры те же, что и в предыдущем примере только для мужчин.

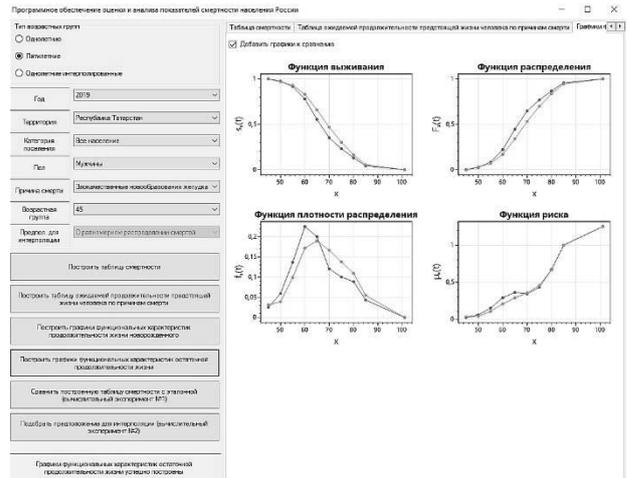


**Рис. 3.** Пример построения графиков функциональных характеристик продолжительности жизни новорожденного



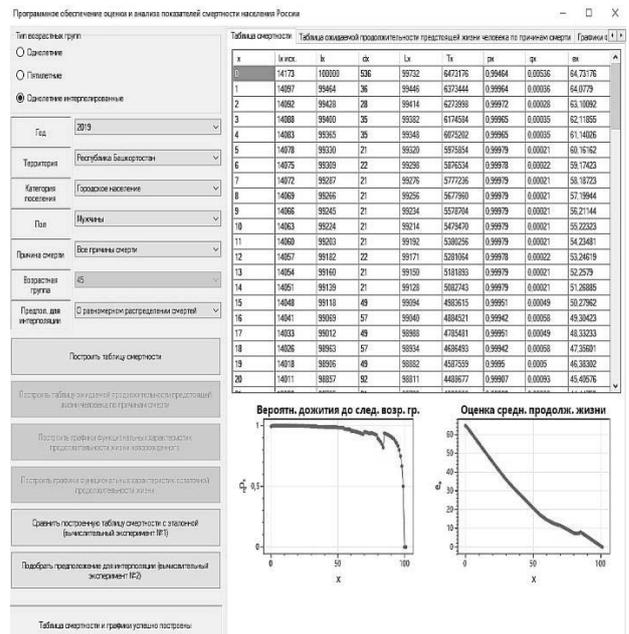
**Рис. 4.** Пример построения графиков функциональных характеристик остаточной продолжительности жизни

При необходимости графики функциональных характеристик продолжительности жизни при изменении одного из входных параметров можно сравнивать друг с другом. На рис. 5 представлено сравнение характеристик для двух регионов – Республики Башкортостан и Республики Татарстан – по одним и тем же значениям остальных входных параметров, заданным в предыдущем примере.



**Рис. 5.** Пример сравнения графиков функциональных характеристик остаточной продолжительности жизни

Как уже было отмечено, что для входных данных, учитывающих причины смерти, имеющихся только для пятилетних возрастных групп, в программе реализовано преобразование краткой таблицы в полную на основе интерполяции. На рис. 6 представлен результат такого преобразования для входных данных из примера 1.



**Рис. 6.** Пример преобразования краткой таблицы смертности в полную

**ЗАКЛЮЧЕНИЕ**

Разработано программное обеспечение обеспечения оценки и анализа показателей выживаемости и смертности. Данное программное

обеспечение работает с Российской базой данных по рождаемости и смертности Центра демографических исследований Российской экономической школы, что дает возможность предоставить пользователю обширную детальную справочную информацию о числовых и функциональных характеристиках продолжительности жизни для различных групп населения России, а также проводить сравнение характеристик при изменении входных параметров.

### БЛАГОДАРНОСТИ

*Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-07-00780.*

### СПИСОК ЛИТЕРАТУРЫ

1. Сайт Президента России. Указ о национальных целях развития России до 2030 года (kremlin.ru). [Электронный ресурс]. URL: <http://kremlin.ru/acts/news/63728> (дата обращения 15.05.2021). [ Site of the President of Russia. Decree on the national development goals of Russia until 2030 (2021, May 15). [Online]. Available: <http://kremlin.ru/acts/news/63728> ]
2. Gardiner J. C. Survival Analysis: Overview of parametric, nonparametric and semiparametric approaches and new developments. [Электронный ресурс]. URL: <https://support.sas.com/resources/papers/proceedings10/252-2010.pdf> (дата обращения 12.12.2020). [ J. C. Gardiner (2020, Dec. 12), *Survival Analysis: Overview of parametric, nonparametric and semiparametric approaches and new developments*. [Online]. Available: <https://support.sas.com/resources/papers/proceedings10/252-2010.pdf> ]
3. Scott W. F. Life Contingencies (Part A2). Edinburg: Herriot-Watt University, 1996. 184 p.
4. Binder G. Construction and Comparison of Mortality Tables Based on Different Techniques. Master Thesis. ETH Zurich, 2014. 132 p.
5. Дарский Л. Е., Тольц М. С. Демографические таблицы. М.: МАКС Пресс, 2013. 104 с. [ L. E. Darskiy, M. S. Tol'ts, *Demographic tables*, (in Russian). Moscow: MAKS Press, 2013. ]
6. Крутько В. Н., Славин М. Б., Смирнова Т. М. Математические основания геронтологии. М.: Едиториал УРСС, 2002. 384 с. [ V. N. Krut'ko, M. B. Slavin, T. M. Smirnova, *Mathematical foundations of gerontology*, (in Russian). Moscow: Editorial URSS, 2002. ]
7. Прокудина Е. И., Абсаттарова Э. Э. Обзор математических моделей и методов анализа данных о выживаемости // Информационные технологии интеллектуальной поддержки принятия решений (ITIDS'2019): VII Всероссийская научная конференция (с приглашением зарубежных ученых), (Уфа, 28–30 мая 2019): тр. конф. Уфа: УГАТУ, 2019. Т. 2. С. 241–244. [ Е. И. Prokudina, Е. Э. Absattarova, "Overview of mathematical models and methods for analyzing survival data", (in Russian), in *Proc. 7<sup>th</sup> All-Russian Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS'2019)*, Ufa, 2019, pp. 241-244. ]

8. Демоскоп Weekly. [Электронный ресурс]. URL: [http://www.demoscope.ru/weekly/ssp/rus\\_ltmenu.php](http://www.demoscope.ru/weekly/ssp/rus_ltmenu.php) (дата обращения 20.03.2021). [ Demoscope Weekly (2021, March. 20). [Online]. Available: <http://www.demoscope.ru/weekly/pril.php> ]

9. Оперативная статистика (bi.gks.ru). [Электронный ресурс]. URL: [http://bi.gks.ru/biportal/contourbi.jsp?allsol=1&solution=Dashboard&project=%2FDashboard%2Fdemography\\_rosstat](http://bi.gks.ru/biportal/contourbi.jsp?allsol=1&solution=Dashboard&project=%2FDashboard%2Fdemography_rosstat) (дата обращения 18.03.2021). [ Operational statistics (2021, Mar. 18). [Online]. Available: [http://bi.gks.ru/biportal/contourbi.jsp?allsol=1&solution=Dashboard&project=%2FDashboard%2Fdemography\\_rosstat](http://bi.gks.ru/biportal/contourbi.jsp?allsol=1&solution=Dashboard&project=%2FDashboard%2Fdemography_rosstat) ]

10. Данные – Европейский портал информации здравоохранения (who.int). [Электронный ресурс]. URL: <https://gateway.euro.who.int/ru/datasets/#hfamdb> (дата обращения 10.05.2021) [ Data - European Health Information Gateway (2021, May 10). [Online]. Available: <https://gateway.euro.who.int/ru/datasets/#hfamdb> ]

11. Российская база данных по рождаемости и смертности. Центр демографических исследований Российской экономической школы. Москва (Россия) (demogr.nes.ru). [Электронный ресурс]. URL: [http://demogr.nes.ru/index.php/ru/demogr\\_indicat/data](http://demogr.nes.ru/index.php/ru/demogr_indicat/data) (дата обращения 5.05.2021). [ The Russian Fertility and Mortality database of the Centre of Demographic Research at the New Economic School (the "Russian Economic School") (2021, May 5). [Online]. Available: [http://demogr.nes.ru/index.php/ru/demogr\\_indicat/data](http://demogr.nes.ru/index.php/ru/demogr_indicat/data) ]

12. Бронштейн Е. М., Прокудина Е. И. Основы актуарной математики. Страхование жизни и пенсионные схемы: учебное пособие. Уфа: УГАТУ, 2002. 132 с. [ Е. М. Bronshteyn, Е. И. Prokudina, *Foundations of actuarial mathematics. Life Insurance and Pension Schemes: A Study Guide*, (in Russian). Ufa: UGATU, 2002. ]

### ОБ АВТОРАХ

**ПРОКУДИНА Елена Ивановна**, доц. каф. вычислительной математики и кибернетики. Дипл. математик (БашГУ, 1984). Канд. физ.-мат. наук по прим. выч. техн., мат. мод. и мат. мет. в науч. иссл. (БашГУ, 1993). Иссл. в обл. математического моделирования, актуарной математики, теории риска.

**ВАХИТОВ Артур Русланович**, инж.-программист ЗАО ЦСИ «Интегро». Дипл. магистр мат. обесп. и админ. информ. систем (УГАТУ, 2021).

### METADATA

**Title:** Estimating and analyzing software of the characteristics of the Russian population survival and mortality.

**Authors:** Е. И. Prokudina <sup>1</sup>, А. Р. Vakhitov <sup>2</sup>

**Affiliation:**

<sup>1</sup> Ufa State Aviation Technical University (UGATU), Russia.

<sup>2</sup> CJSC System Research Center «Integro», Russia.

**Email:** <sup>1</sup>preliv@gmail.com, <sup>2</sup>ejwhfkw@gmail.com

**Language:** Russian.

**Source:** SIIT (scientific journal of Ufa State Aviation Technical University), vol. 4, no. 1 (8), pp. 44-50, 2022. ISSN 2686-7044 (Online), ISSN 2658-5014 (Print).

**Abstract:** The problem of estimating and analyzing the numerical and functional characteristics of human future lifetime

for the population of Russia, which is differentiated according to various criteria: age, sex, region, type of settlement, cause of death, is considered. Software has been developed based on algorithms for calculating the characteristics of lifetime by mortality rates from the Russian Fertility and Mortality database of the Centre of Demographic Research at the New Economic School (the "Russian Economic School"). The software product implements the construction of life tables, visualization of the functional characteristics of life time, and their comparative analysis.

**Key words:** algorithms for estimating survival and mortality rates; software; demographic database; population of Russia.

**About authors:**

**PROKUDINA, Elena Ivanovna**, Assoc. Prof., Dept. of Computational Mathematics and Cybernetics. Dipl. Mathematics (BashSU, 1984). Cand. of Phis.-Math. Sci. (BashSU, 1993).

**VAKHITOV, Artur Ruslanovich**, Software engineer, CJSC System Research Center "Integro". Dipl. Master of Software and Administration of Information System (USATU, 2021).