

Математическое моделирование социально-демографических процессов на примере статистических показателей населения РФ

Хорольская Ирина Витальевна

Тихоокеанский государственный медицинский университет

старший преподаватель

Аннотация

В статье дается описание матричной модели. В работе используется матричная модель для описания и прогнозирования характеристик возрастно-половой структуры населения Российской Федерации, сделан прогноз на 2016-2018 гг. с применением статистических показателей.

Ключевые слова: математическая модель, прогнозирование, демографические процессы, популяция

Mathematical modeling of social-demographic processes on the example of statistical indices of the Russian population

Horol'skaya Irina Vital'evna

Pacific National Medical University

senior lecturer

Abstract

The article describes the matrix model. The work uses a matrix model to describe and predict the characteristics of the age and sex structure of the population of the Russian Federation, a forecast for 2016-2018 is made. with the use of statistical indicators.

Keywords: mathematical model, forecasting, demographic processes, population

Введение

Метод матричного анализа представляет собой исследование взаимосвязей между различными объектами с помощью матричных моделей. Матричные модели строятся в виде таблиц.

Представление информации с помощью матричных моделей можно считать основополагающим для социометрии, для различных классов экономико-математических моделей, например, приводимых к задачам дискретного, линейного и нелинейного программирования. Кроме этого, оно широко используется в математической статистике, теории графов, теории игр [1].

Матричная модель может рассматриваться как вариант динамической модели. Один из ранних аналогов был разработан Лесли и Льюисом как модель, прогнозирующая будущую возрастную структуру популяции по предполагаемым коэффициентам плодovitости и выживания и известной

возрастной структуре в данный момент времени. Предполагается, что число возрастных групп в популяции равно n .

Пусть функции $S_i(x_1 \dots x_n)$, которые описывают переход из i -той возрастной группы в $(i+1)$ -ю группу, являются линейными функциями численности i -той возрастной группы, $i = 1, 2, \dots, n-1$, где S_i ($0 < S_i \leq 1$) – коэффициенты выживаемости, показывающие, какая доля особей i -той группы доживет до $(i+1)$ -го возраста. Тогда для всех групп, начиная со второй, будет выполняться:

$$X_{i+1}(t+1) = S_i * X_i(t), i = 1, 2, 3, \dots, n-1 \quad (1)$$

Вектор-столбец, координатами которого являются численности всех возрастных групп, обозначим через $X(t)$. Тогда:

$$X(t+1) = L * X(t), \quad (2)$$

$$\text{где } L = \begin{vmatrix} b_1 & \dots & b_{n-1} & b_n \\ S_1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & S_{n-1} & 0 \end{vmatrix} \quad (3)$$

где $b_i \geq 0, i = 1 \dots n$;

$0 < S_j \leq 1, j = 1, 2 \dots n$;

L – квадратная матрица, имеющая порядок $n * n$ и называемая матрицей Лесли [2].

Уравнение (2) с матрицей (3) имеет название модель Лесли. Стоит упомянуть о ряде сложностей, которые появляются в приложениях модели Лесли к популяциям и связаны с ограничениями модели. Допустим, не удастся рассмотреть в n -ной (последней) возрастной группе только особей последнего репродуктивного возраста. Тогда к n -ной группе относят всех старших особей, а к матрице Лесли добавляют элемент p_n ($0 < p_n < 1$), представляющий долю особей n -ной группы, выживающих за один временной интервал. Тогда матрицу L можно представить следующим образом:

$$L + \Delta(p_n) = \begin{vmatrix} b_1 & \dots & b_{n-1} & b_n \\ S_1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & S_{n-1} & p_n \end{vmatrix} \quad (4)$$

$$X_n(t+1) = S_{n-1} * X_{n-1}(t) + p_n * X_n(t) \quad (5)$$

В этом случае получается, что некоторая часть популяции живет бесконечно долго.

Другая сложность заключается в том, что не всегда можно выбрать масштаб времени так, что временные моменты $t=1, 2, \dots$ соответствуют переходам из одной возрастной группы в другую. Тогда поступают следующим образом: вместе с величинами S_i рассматривают величины p_i ($0 < p_i < 1$), представляющие долю тех особей i -той группы, которые не перешли в следующую возрастную группу к моменту времени i . Тогда матрица L имеет вид:

$$L + \Delta(p_1 \dots p_n) = \begin{vmatrix} b_1 + p_1 & b_2 & \dots & b_n \\ S_1 & p_2 & \dots & 0 \\ 0 & \dots & S_{n-1} & p_n \end{vmatrix} \quad (6)$$

Особенности демографических моделей:

1. невозможность проведения активных экспериментов над человеческой популяцией для более точного построения моделей;
2. изменения параметров модели (коэффициент рождаемости, коэффициент смертности и т.д.) определяются, помимо биологических возможностей, различными социально-экономическими, политическими, психологическими факторами, которые являются трудно предсказуемыми;
3. данные берутся на основании выборочных наблюдений или статистических показателей, что приводит к большим ошибкам по сравнению с моделями, применяемыми, например, в технических системах;
4. «плавное» изменение численности, что затрудняет решение задачи идентификации модели на ограниченном отрезке времени [3].

В данной статье основное внимание будет сконцентрировано на моделировании естественного движения населения.

Оценка и прогнозирование численности и структурного состава населения регионов РФ является в настоящее время актуальной задачей.

Цель данной статьи заключается в решении задачи текущего оценивания и прогнозирования численности и возрастно-полового состава населения РФ.

Вычислительные эксперименты

Для описания демографической модели разобьем население РФ на группы в соответствии с возрастом.

Рассмотрим первый случай. Разобьем население РФ на группы в соответствии с возрастом, не учитывая пол. Величину временного интервала, охватывающего одну возрастную группу, положим равной 10 годам. Следовательно, получим, что население РФ состоит из 9 возрастных групп: 0-9 лет, 10-19 лет, 20-29 лет, 30-39 лет, 40-49 лет, 50-59 лет, 60-69 лет, 70-79 лет, 80 лет и старше.

Для осуществления оценки и прогнозирования численности и структурного состава населения РФ используем официальные данные Росстата (используем данные за 2013-2017 года) [4]. Используя данные статистики за 2015 год, составим вектор-столбец $X(t)$, координатами которого являются численности всех возрастных групп. Поставим следующую задачу: пользуясь теоретическими источниками и данными Росстата, попытаемся осуществить прогнозирование численности населения на 2016 год. Для этого составим матрицу L и найдем вектор $X(t+1) = L * X(t)$, координатами которого будут значения прогнозируемой численности возрастных групп на 2016 год. Находим неизвестные элементы матрицы L :

$$b_1=0, b_2=0,013, b_3=0,05, b_4=0,0296, b_5=0,002, b_6=b_7=b_8=b_9=0;$$

$$p_1=p_4=0,9, p_2=p_3=p_5=p_8=0,89, p_6=0,894, p_7=0,94, p_9=0,82$$

$S_1=S_4=S_6=0,09, S_2=S_3=S_5=0,1, S_7=0,04, S_8=0,0645$, где b_i - коэффициенты рождаемости в соответствующих возрастных группах, p_i - величины, означающие долю особей из группы, не успевших перейти в следующий возрастной класс, $S_i (0 < S_i \leq 1)$ - коэффициенты выживаемости, показывающие, какая доля особей i -той группы доживет до $(i+1)$ -го возраста.

При расчете данных коэффициентов были составлены соотношения, при их составлении использовались официальные статистические данные [4]. Наличие в теоретических источниках данных по соответствующим возрастным группам позволило составить соотношения и рассчитать необходимые коэффициенты за 2013-2015 гг. В качестве элементов матрицы использовались средние значения соответствующих коэффициентов за указанные года.

Найденные координаты вектора $X(t+1)$ незначительно отличаются от официальных статистических показателей за 2016 год. Общая численность населения РФ по данным статистики составляет 146545000 человек [4]. Прогнозируемая численность населения составила 143164221 чел. Погрешность вычисления составляет 2,3 %. Гистограмма официальной и прогнозируемой численности населения на 2016 год представлена на рис. 1.

Аналогичным образом, используя составленную матрицу L и полученный путем прогнозирования вектор с координатами численности возрастных групп на 2016 год, по формуле (2) можно сделать прогноз на 2017 год. Используя данный метод, удалось осуществить прогнозирование численности и возрастного состава населения РФ на 2016-2018 гг. Полученные результаты сравнили с официальными статистическими данными за 2016-2017 года (Таблица 1). Также в данной таблице представлены полученные путем прогнозирования данные численности населения за 2018 г. Ошибка вычислений составила 3 % для 2017 года. Соответствующая гистограмма представлена на рис. 2.

Таблица 1- Численность и возрастной состав населения РФ на 2016-2018 гг.

Возраст	2016		2017		2018
	прогноз	оф.д.	прогноз	оф.д.	прогноз
0-9	17267532	17730000	17453896	18140 000	17571915
10-19	13772244	13985000	13811375	14098000	13862974
20-29	20590756	20857000	19702997	19707000	18916805
30-39	22523723	23317000	22330426	23731000	22067683
40-49	18919214	19413000	18865236	19661000	18799798
50-59	21105613	21449000	20760339	20990000	20446267
60-69	16422599	16708000	17336748	17247000	18164974
70-79	8427664	8905000	8157525	8862000	7953667
80 и более	4134876	4181000	3934183	4368000	3752190
Всего	143164221	146545000	142352725	146804000	141536273

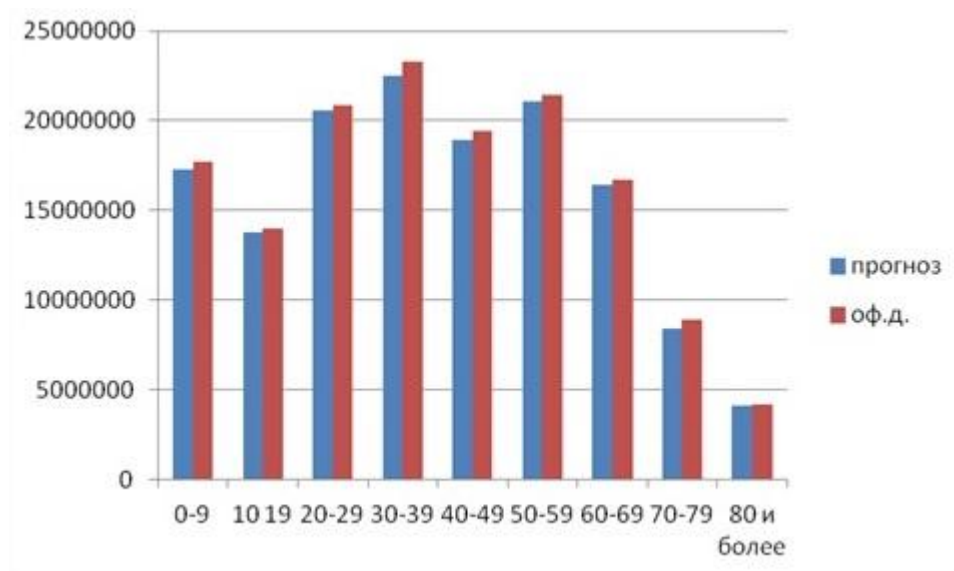


Рисунок 1 - Показатели численности населения РФ на 2016 год

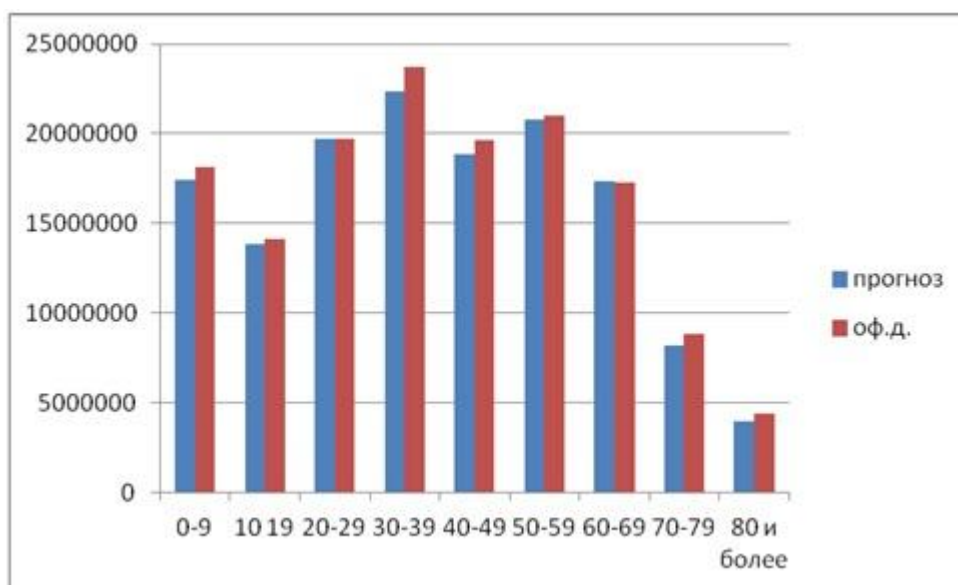


Рисунок 2 - Показатели численности населения РФ на 2017 год

Была оценена значимость различий официальных статистических данных и прогнозируемых данных за 2016-2017 гг. по критерию Стьюдента.

t -критерий Стьюдента был рассчитан по формуле:

$$t = \frac{|M_1 - M_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (7)$$

где M_1 , M_2 – средние арифметические групп, σ_1 , σ_2 – стандартные отклонения групп, n_1 , n_2 – кол-во значений в группах. Полученные данные представлены в Таблице 2.

Таблица 2 - Оценка значимости различий официальных статистических данных и прогнозируемых данных за 2016-2017 гг. по критерию Стьюдента

	2016 г	2017 г
M_1	15907136	15816969
M_2	16282778	16311556
σ_1	6154420	6125097
σ_2	6278252	6194214
t	0,128	0,17

Как видно из табличных данных, значения t составили: для 2016 г. - 0,128; для 2017 г. - 0,17. Число степеней свободы в нашем случае равно 16.

Сравнив рассчитанные значения t -критерия с табличными значениями при соответствующей степени свободы, можно сделать вывод, что различия между официальными данными и прогнозом на 2016 - 2017 гг. не являются значимыми (рассчитанные значения t -критерия Стьюдента меньше табличных значений).

Рассмотрим второй случай. Население РФ разобьем на группы в соответствии с возрастом, но при этом учтем половой состав населения. Тогда матрица Лесли будет иметь вид:

$$\mathbf{M} = \left(\begin{array}{ccc|ccc}
 p_1^m & \dots & 0 & b_1^m & \dots & b_n^m \\
 \dots & \dots & \dots & \dots & \dots & \dots \\
 0 & \dots & S_{n-1}^m p_n^m & 0 & \dots & 0 \\
 \dots & \dots & \dots & b_1^w + p_1^w & \dots & b_n^w \\
 \dots & \dots & \dots & \dots & \dots & \dots \\
 0 & \dots & 0 & 0 & \dots & S_{n-1}^w p_n^w
 \end{array} \right)$$

Находим неизвестные элементы матрицы \mathbf{M} :

$$b_1^m=0, b_2^m=0,00665, b_3^m=0,025, b_4^m=0,0148, b_5^m=0,001, b_6^m=0, b_7^m=0, b_8^m=0, b_9^m=0.$$

$$b_1^w=0, b_2^w=0,00665, b_3^w=0,025, b_4^w=0,0148, b_5^w=0,001, b_6^w=0, b_7^w=0, b_8^w=0, b_9^w=0.$$

Коэффициенты p_i для мужчин равны соответственно 0,9; 0,89; 0,89; 0,9; 0,89; 0,894; 0,93; 0,87; 0,82, для женщин – 0,9; 0,89; 0,89; 0,9; 0,89; 0,9; 0,945; 0,895; 0,82

Коэффициенты S_i для мужчин равны соответственно 0,09; 0,1; 0,1; 0,09; 0,1; 0,09; 0,04; 0,06, для женщин – 0,09; 0,1; 0,1; 0,09; 0,1; 0,097; 0,04; 0,066.

При расчете данных коэффициентов были составлены соотношения, при их составлении использовались официальные статистические данные [4]. Аналогично первому случаю, наличие в теоретических источниках данных по соответствующим возрастным группам позволило составить соотношения и рассчитать необходимые коэффициенты за 2013-2015 гг. В качестве элементов матрицы использовались средние значения соответствующих коэффициентов за указанные года.

Продлав аналогичные вычисления, получим прогноз численности и структурного состава населения на 2016-2018 гг. Показатели численности мужского и женского населения РФ на 2016-2017 гг. соответственно представлены в таблицах 3 и 4. Прогнозируемая численность населения на 2016 год составляет 142286106 человек, на 2017 год – 140628606 человек. Погрешность вычисления составляет 2,9% на 2016 год, 4,2% на 2017 год.

Таблица 3 - Показатели численности мужского населения РФ на 2016-2017 гг.

Возраст	2016		2017	
	прогноз	оф.д.	прогноз	оф.д.
0-9	8337761	9100000	7979801	9313000
10- 19	7054188	7156000	7028626	7209000
20-29	10457381	10596000	10012488	10028000
30-39	11171535	11548000	11100120	11775000
40-49	9124869	9326000	9126572	9447000
50-59	9565015	9681000	9463610	9492000
60-69	6622797	6688000	7020053	6906000
70-79	2656860	2710000	2576380	2952000
80 и более	1022402	1092000	997781	922000
Всего	66012808	67897000	65305431	68044000

Таблица 4 - Показатели численности женского населения РФ на 2016-2017 гг.

Возраст	2016		2017	
	прогноз	оф.д.	прогноз	оф.д.
0-9	7942671	8630000	7624220	8827000
10- 19	6718056	6829000	6693910	6889000
20-29	10133375	10261000	9690509	9679000
30-39	11352188	11769000	11230307	11956000
40-49	9794345	10087000	9738664	10214000
50-59	11611441	11768000	11429731	11498000
60-69	9866752	10020000	10450390	10341000
70-79	5745449	6195000	5536847	5910000
80 и более	3109021	3089000	2928597	3446000
Всего	76273298	78648000	75323175	78760000

Полученные путем прогнозирования данные численности населения за 2018 г. представлены в Таблице 5.

Таблица 5 – Прогноз численности населения на 2018 г.

Возраст	Мужчины	Женщины
0-9	7644545	7324522
10- 19	6973659	6643760
20-29	9613977	9293944

30-39	10991357	11076327
40-49	9121660	9678139
50-59	9373125	11260624
60-69	7380374	10984302
70-79	2522253	5373494
80 и более	972763	2766881
Всего	64593713	74401993

Аналогичным образом была оценена значимость различий официальных статистических данных и прогнозируемых данных за 2016-2017 гг. по критерию Стьюдента. Полученные данные представлены в Таблице 6.

Таблица 6 - Оценка значимости различий официальных статистических данных и прогнозируемых данных за 2016-2017 гг. по критерию Стьюдента.

	2016 г		2017 г	
	Мужчины	Женщины	Мужчины	Женщины
M_1	7334756	8474811	7256159	8369242
M_2	7544111	8738667	7560444	8751111
σ_1	3463567	2831910	3399328	2882911
σ_2	3560009	2876298	3536794	2807018
t	0,126	0,196	0,186	0,285

Как видно из табличных данных, значения t составили: для 2016 г. – 0,126 для мужчин и 0,196 для женщин; для 2017 г. – 0,186 для мужчин и 0,285 для женщин. Число степеней свободы равно 16.

Сравнив рассчитанные значения t -критерия с табличными значениями при соответствующей степени свободы, можно сделать вывод, что различия между официальными данными и прогнозом на 2016 - 2017 гг. не являются значимыми (рассчитанные значения t -критерия Стьюдента меньше табличных значений для всех уровней значимости).

Заключение

Проведенные вычислительные эксперименты подтверждают возможность решения задачи текущего оценивания и прогнозирования численности и возрастно-полового состава населения.

С помощью методов математической статистики доказано, что различия между официальными показателями численности и структурного состава населения РФ и прогностическими данными, полученными при использовании модели Лесли, не являются значимыми. Представленная матричная модель позволяет осуществить оценку и прогнозирование численности и структурного состава населения.

Библиографический список

1. Власов М. П., Шимко П. Д. Моделирование экономических процессов. Ростов-на-Дону: Феникс, 2005.
2. Логофет Д. О., Белова И. Н. Неотрицательные матрицы как инструмент моделирования динамики популяций: классические модели и современные обобщения // Фундаментальная и прикладная математика. 2007. № 4. С. 145-164.
3. Рудалев В. Г., Кремер А. И. Математическая модель динамики структурного состава населения // Вестник ВГУ. 2007. №1.
4. Демографический ежегодник России. Стат. сборник. М., 2014-2017.
5. Абакумов А. И., Гиричева Е. Е. Моделирование зависимости возрастной структуры населения региона от его доходов // Дальневосточный математический журнал. 2003. №1. С. 316-325.
6. Абакумов А.И., Гиричева Е.Е. Возрастная структура населения в зависимости от доходов // Обозрение прикладной и промышленной математики. 2006. № 1. С. 51-64.
7. Камалян Р. З., Камалян С. Р., Гергенридер К. А. Сравнительный анализ моделей роста населения Земли // Вестник ИМСИТа. 2007. №1,2.
8. Кирьянов Д. А. Прогноз рождаемости при моделировании роста численности населения // Вестник Пермского университета. 2016. № 3.