

DOI: 10.12731/2218-7405-2017-6-56-66

УДК 31, 519.23

ПОСТРОЕНИЕ СТАТИСТИЧЕСКОЙ МОДЕЛИ ОБЩЕЙ ЧИСЛЕННОСТИ НАСЕЛЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ НА ОСНОВЕ РЕТРОСПЕКТИВНОГО ПРОГНОЗА

Кочегарова О.С., Лажануинкас Ю.В.

В статье рассмотрен ретроспективный прогноз общей численности населения Российской Федерации за период 2001–2017 гг. Проведен сравнительный анализ фактического значения показателя общей численности населения РФ на 20.03.2017 г. по данным Федеральной службы государственной статистики РФ с полученным прогнозным значением. Выбор модели прогнозирования осуществлялся методом подбора кривых роста на основе корреляционно-регрессионного анализа и методом наименьших квадратов. Качественный подбор уравнения регрессии определялся наименьшей ошибкой аппроксимации уровней временного ряда. Анализ значимости подобранного уравнения регрессии статистическими методами позволяет сделать вывод о верном подборе модели и возможности ее использования для прогноза численности населения.

Цель: подбор уравнения регрессии для изучения динамики численности населения на основе эмпирических данных и оценка его точности для прогноза численности населения.

Метод или методология проведения работы: метод подбора кривых роста на основе корреляционно-регрессионного анализа и метод наименьших квадратов.

Результаты: получено подтверждение эффективности построенной модели для прогнозов демографических процессов.

Область применения результатов: полученные результаты целесообразно применять при построении прогнозов демографических процессов.

Ключевые слова: демография; численность населения; ретроспективный прогноз; статистическая модель; точность прогноза.

CONSTRUCTION OF STATISTICAL MODEL THE OVERALL POPULATION OF THE RUSSIAN FEDERATION ON THE BASIS OF RETROSPECTIVE FORECAST

Kochegarova O.S., Lazhauninkas J.V.

The article considers the retrospective forecast of the total population of the Russian Federation for the period 2001–2017. comparative analysis of the actual values of the total population of the Russian Federation on 20.03.2017 according to the Federal state statistics service of the Russian Federation received the forecast value. Model selection forecasting was carried out by the method of selection of growth curves on the basis of correlation and regression analysis and least squares method. A quality selection of the regression equation was determined with the least error of approximation of time series levels. Analysis of the significance of the selected regression equation by statistical methods allows to make a conclusion about the right choice of models and the possibility of its use for population estimates.

Purpose: *to estimate the significance of selected regression equations for the forecast of the population.*

Methodology in article: *the fitting of growth curves on the basis of correlation and regression analysis and least squares method.*

Results: *received confirmation of the effectiveness of the constructed model for forecasts of demographic processes.*

Practical implications: *the obtained results should be used when building forecasts of demographic processes.*

Keywords: *demographics; population; retrospective forecast; statistical model; forecast accuracy.*

Для проведения эффективной социально-экономической политики страны необходимо обладать данными о численности населения. Это помогает заглянуть в демографическое будущее и предсказать возможные демографические сдвиги. Повышенный интерес государства к данной проблеме нашел свое отражение в Указе Прези-

дента Российской Федерации об утверждении Концепции демографической политики Российской Федерации на период до 2025 года № 1351 от 09.10.2007 г. [1].

Задача нашего исследования состояла в построении ретроспективного прогноза для периода 2001–2017 гг., за который уже имеются фактические значения, с целью проверки полученной модели и возможности ее использования для дальнейших прогнозов.

Математические модели являются наиболее распространенными методами прогнозирования. В зависимости от характера исходной информации и поставленной цели мы вправе использовать различные методы прогнозирования [2, 3, 4, 5, 8].

Для проверки ряда на наличие нелинейной тенденции были вычислены линейные коэффициенты автокорреляции для временного ряда, состоящего из логарифмов исходных уровней. Отличные от нуля значения коэффициентов автокорреляции свидетельствовали о наличии нелинейной тенденции для рассматриваемого признака.

Изменение численности населения подчиняется закону геометрической прогрессии, поэтому наибольшее распространение среди математических формул для прогнозирования численности населения получила показательная функция и, в частности, экспоненциальная. Согласно этому численность населения, прогнозируемая по истечении x лет, рассчитывается по экспоненциальной модели:

$$Y_x = e^{a+b \cdot x}, \quad (1)$$

где Y_x – прогнозное значение, a и b – коэффициенты уравнения регрессии.

Подбор коэффициентов уравнений регрессий проводился методом наименьших квадратов по известным формулам:

$$b = \frac{n \cdot \sum x_i \cdot \ln y_i - \sum x_i \cdot \sum \ln y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2},$$
$$a = \frac{1}{n} \cdot \sum \ln y_i - \frac{b}{n} \cdot \sum x_i.$$

Для анализа точности количественных прогнозов используют абсолютные, относительные и сравнительные показатели. Практическую значимость представляют

$$ME = \frac{\sum_{t=1}^n (y_i - y_t)}{n} - \text{средняя абсолютная ошибка,} \quad (2)$$

$$MAPE = \frac{1}{n} \cdot \sum_{t=1}^n \frac{|y_i - y_t|}{y_i} \cdot 100\% - \text{ошибка аппроксимации.} \quad (3)$$

Качество прогноза определяется его точностью. При этом в самом широком понимании точность прогноза определяется по формуле:

$$T_{\text{прогноза}} = (1 - ME) \cdot 100\%. \quad (4)$$

Однако точное совпадение фактических данных с прогнозируемым значением маловероятно. Отсюда следует, что любой прогноз носит приближенный характер. Поэтому для устранения ошибки от точечной оценки прогноза следует перейти к интервальной оценке или указать, так называемый, доверительный интервал. Нелинейность прогностической модели, скорее всего, подразумевает, что распределение прогнозируемых значений несимметричное, что, в свою очередь, подразумевает несимметричность доверительного интервала. При нормальном законе распределения случайной величины доверительный интервал линейного уравнения регрессии может быть построен традиционным способом с использованием *t*-распределения Стьюдента. Но при нелинейной регрессии данный метод не учитывает ряд особенностей эмпирического распределения данных, например, его асимметрию. В этом случае необходимо перейти от нелинейной регрессионной модели к линейной с помощью линеаризирующего или нормализующего преобразований, построить доверительный интервал для линейной регрессии, а затем с помощью обратного преобразования, например, преобразования Джонсона, построить доверительный интервал для нелинейной модели. Но в цели данной работы построение доверительного интервала не входило и было оставлено для дальнейших исследований [12, 13].

Для решения поставленной задачи были взяты официальные данные Федеральной службы государственной статистики по об-

щей численности населения в период 2001–2017 гг. [14]. Заметим, что данные до 2015 года приведены без учета населения Крыма и Севастополя.

Для упрощения расчета коэффициентов уравнения регрессии номера уровней ряда будем считать как $x=1, 2, 3, 4, 5, 6, \dots, 17$, то есть: 1 – 2001 год, 2 – 2002 год и так далее.

Для упрощения вычислений коэффициентов составим расчетную таблицу 1 в среде табличного процессора MS Excel [9]:

Таблица 1.

Расчет коэффициентов уравнения регрессии

Номер года, x_i	Год	Фактическое значение, млн чел, y_i	x_i^2	$\ln y_i$	$x_i \cdot \ln y_i$
1	2001	146,3	1,0	4,99	4,99
2	2002	145,2	4,0	4,98	9,96
3	2003	145,0	9,0	4,98	14,93
4	2004	144,3	16,0	4,97	19,89
5	2005	143,8	25,0	4,97	24,84
6	2006	143,2	36,0	4,96	29,79
7	2007	142,8	49,0	4,96	34,73
8	2008	142,8	64,0	4,96	39,69
9	2009	142,7	81,0	4,96	44,65
10	2010	142,9	100,0	4,96	49,62
11	2011	142,9	121,0	4,96	54,58
12	2012	143,0	144,0	4,96	59,55
13	2013	143,3	169,0	4,96	64,54
14	2014	143,7	196,0	4,97	69,55
15	2015	146,3	225,0	4,99	74,78
16	2016	146,5	256,0	4,99	79,79
17	2017	146,8	289,0	4,99	84,81
	сумма		1 785	84,5103	760,6987

Тогда, используя данные таблицы 1, вычислим параметры регрессии:

$$b = \frac{17 \cdot 760,6987 - 153 \cdot 84,5103}{17 \cdot 1785 - 153^2} \approx 0,0003,$$

$$a = \sum \frac{1}{17} \cdot 84,5103 - \frac{0,0003}{17} \cdot 153 \approx 4,9688.$$

Подставив найденные параметры a и b в формулу (1) получаем уравнение регрессии, которое описывает математическую модель динамики численности населения:

$$Y_x = e^{4,9688+0,0003 \cdot x}$$

По данному уравнению вычислим ретроспективные прогнозы численности населения за период 2001–2017 гг. Произведем расчет в MS Excel. Получим вспомогательную таблицу 2:

Таблица 2.

Построение ретроспективного прогноза и оценка его качества

Номер года, x_i	Год	Фактическое значение, млн чел. y_i	Прогнозное значение, млн чел. y_i	$y_i - y_i$	$\frac{ y_i - y_i }{y_i}$
1	2001	146,3	143,90	2,4	0,0164
2	2002	145,2	143,94	1,26	0,0087
3	2003	145,0	143,98	1,02	0,0070
4	2004	144,3	144,03	0,27	0,0019
5	2005	143,8	144,07	-0,27	0,0019
6	2006	143,2	144,11	-0,91	0,0064
7	2007	142,8	144,16	-1,36	0,0095
8	2008	142,8	144,20	-1,4	0,0098
9	2009	142,7	144,24	-1,54	0,0108
10	2010	142,9	144,29	-1,39	0,0097
11	2011	142,9	144,33	-1,43	0,0100
12	2012	143,0	144,37	-1,37	0,0096
13	2013	143,3	144,42	-1,12	0,0078
14	2014	143,7	144,46	-0,76	0,0053
15	2015	146,3	144,50	1,8	0,0123
16	2016	146,5	144,55	1,95	0,0133
17	2017	146,8	144,59	2,21	0,0151
	сумма	2 451,5	2 452,14	-0,64	0,1554

Правильность расчета параметров уравнения регрессии может быть проверена с помощью сравнения сумм $\sum y_i \approx \sum y_i = 2452$. В данном случае получено некоторое расхождение сумм, но их разница никак не влияет на дальнейшие расчеты.

По формулам (2), (3) и данным расчетной таблицы 2 получим

$$ME = \frac{-0,64}{17} = -0,038, \quad MAPE = \frac{1}{17} \cdot 0,1554 \cdot 100\% = 0,91\%.$$

Вычисленное значение $ME = -0,038$ позволяет определить точность по формуле (4): $T_{\text{прогноза}} = (1 - (-0,038)) \cdot 100\% = 103,8\%$.

Итак, нами реализован способ уточнения математического аппарата для статистического анализа общей численности населения РФ. Построено уравнение регрессии, которое описывает математическую модель динамики численности населения на основе данных с 2001 по 2017 год. По полученному уравнению сделан ретроспективный прогноз на указанный период. Для анализа точности прогноза вычислены показатели, представляющие практическую значимость. Точность прогноза составляет 103,8%. Ошибка аппроксимации подбора уравнения регрессии получилась равной $MAPE = 0,91\% < 10\%$, что говорит о высокой точности прогноза. Таким образом, модель может быть использована для прогнозирования демографических процессов.

Результаты исследования рекомендуется использовать в процессе изучения демографических процессов, что помогает заглянуть в демографическое будущее и предсказать возможные демографические сдвиги.

Проведенное исследование не исчерпывает всей полноты рассматриваемой проблемы, а предлагает один из путей ее решения. Более глубокого изучения, на наш взгляд, требуют вопросы построения доверительного интервала для полученного уравнения нелинейной (экспоненциальной) регрессии для построения демографических прогнозов и оценки их точности.

Список литературы

1. Указ Президента Российской Федерации № 1351 от 09.10.2007 г. «Об утверждении Концепции демографической политики Российской Федерации на период до 2025 года» // Собрание законодательства РФ. 2007. № 42. Ст. 5009.
2. Абдувасиева З.С. О некоторых математических моделях численности населения республики Таджикистан // Вестник университета. 2014. Т. 1, №4 (47). С. 120–126.
3. Азаров С.В. Регрессионные методы статистического оценивания в социальных исследованиях / С.В. Азаров, Е.В. Черепанов // Математические методы и компьютерные технологии в маркетинговых и социальных исследованиях. Сборник научных работ. М.: Академия менеджмента инноваций, 2004. №4. С. 56–72.

4. Гришин А.Ф., Кочерова Е.В. Статистические модели: построение, оценка, анализ: учеб. пособие. М.: Финансы и статистика, 2005. 417 с.
5. Гулиян Б.Ш. О вопросах математического моделирования и прогноза численности населения региона на примере РФ / Б.Ш. Гулиян, Г.Б. Гулиян, И.А. Нестеров // Славянский форум. 2016. № 3(13). С. 61–66.
6. Зозуля П.В. Прогнозы численности населения России и мира / П.В. Зозуля, А.В. Зозуля // Вестник Университета (Государственный университет управления). 2017. №2 (февраль). С. 71–75.
7. Кочегарова О.С. ABC-анализ как средство оптимизации решения задач прикладной статистики / О.С. Кочегарова, Ю.В. Лажауникас // Наука Красноярья. 2016. № 3–3 (26). С. 101–109.
8. Кочегарова О.С. Прогнозирование экономических процессов методами корреляционно-регрессионного анализа / О.С. Кочегарова, Ю.В. Лажауникас // О вопросах и проблемах современных математических и естественных наук: сб. науч. тр. Челябинск, 2015. С. 9–12.
9. Лажауникас Ю.В. Методика проведения статистической обработки экспериментальных данных с использованием табличного процессора Excel / Ю.В. Лажауникас, О.С. Кочегарова // Фундаментальные и прикладные исследования в высшей аграрной школе. Под ред. М.В. Муравьевой и Г.Н. Камышовой. Саратов, 2014. С. 34–36.
10. Мукаева Л.А. Ретроспективный анализ динамики численности населения Чеченской Республики / Л.А. Мукаева, Л.Ш. Батыжева, Л.Т. Солтахмадова // Вестник Чеченского государственного университета. 2015. №1. С. 183–186.
11. Поляков Е.М. Демографическая ситуация в современной России // Общество: социология, психология, педагогика. 2012. № 2. С. 17–21.
12. Приходько С.Б. Доверительный интервал нелинейной регрессии времени восстановления работоспособности устройств терминальной сети / С.Б. Приходько, Л.Н. Макарова // Восточно-Европейский журнал передовых технологий. Харьков, 2014. № 3/4(69). С. 26–29.
13. Приходько С.Б. Інтервальне оцінювання статистичних моментів негаусівських випадкових величин на основі нормалізуючих перетворень // Математичне моделювання: Науковий журнал. Дніпродзержинськ: ДДТУ. 2011. №1(24). С. 9–13.

14. Федеральная служба государственной статистики. <http://www.gks.ru> (дата обращения 24.05.2017).
15. Черепанов Е.В. Вероятностно-статистические основы прикладной социологии и маркетинговых исследований / Акад. Менеджмента Инноваций, Каф. «Аналитический менеджмент», Бюро экономических и социальных технологий-БЭСТ АМИ. Москва, 2006.
16. Якубовский-Липский Ю.О. Изменение компонентов общей численности населения на территории российской федерации с 1745 года по 2010 год // Экологическая экспертиза. 2013. №2. С. 3–26.

References

1. *Sobranie zakonodatel'stva RF. 2007. № 42. St. 5009* [Collection of the legislation of the Russian Federation. 2007. № 42. Art. 5009].
2. Abduvasieva Z.S. O nekotorykh matematicheskikh modelyakh chislenosti naseleniya respubliki Tadjikistan [On some mathematical models of population of the Republic of Tajikistan]. *Vestnik universiteta*. 2014. Vol. 1, no. 4(47), pp. 120–126.
3. Azarov S.V. Regressionnyye metody statisticheskogo otsenivaniya v sotsial'nykh issledovaniyakh [Regression methods of statistical evaluation in social studies]. *Matematicheskie metody i komp'yuternye tekhnologii v marketingovykh i sotsial'nykh issledovaniyakh. Sbornik nauchnykh rabot* [Mathematical methods and computer technologies in marketing and social research. Collection of scientific works]. M.: Akademiya menedzhmenta innovatsii. 2004. no. 4, pp. 56–72.
4. Grishin A.F., Kocherova E.V. *Statisticheskie modeli: postroenie, otsenka, analiz* [Statistical models: construction, estimation, analysis]. Moscow: Finansy i statistika, 2005. 417 p.
5. Guliyani B.Sh. O voprosakh matematicheskogo modelirovaniya i prognoza chislenosti naseleniya regiona na primere RF [On the mathematical modeling and forecast of the population of the region on the example of the Russian Federation]. *Slavyanskii forum*. 2016. no. 3(13), pp. 61–66.
6. Zozulya P.V. Prognozy chislenosti naseleniya Rossii i mira [Projections of the population of Russia and the world]. *Vestnik Universiteta (Gosudarstvennyi universitet upravleniya)*. 2017. №2 (fevral'), pp. 71–75.
7. Kochegarova O.S. ABS-analiz kak sredstvo optimizatsii resheniya zadach prikladnoy statistiki [ABC-analysis as a means to optimize the solu-

- tion of tasks applied statistics]. *Nauka Krasnojars'ja*. 2016. no. 3-3 (26), pp. 101–109.
8. Kochegarova O.S. *Prognozirovanie ekonomicheskikh processov metodami korrelyacionno-regressionnogo analiza* [Forecasting of economic processes by methods of correlation and regression analysis]. Chelyabinsk, 2015, pp. 9–12.
 9. Lazhauninkas J.V. *Metodika provedeniya statisticheskoy obrabotki jeksperimental'nyh dannyh s ispol'zovaniem tablichnogo processora Excel* [The methodology of statistical processing of experimental data using the Excel spreadsheet processor]. Saratov, 2014, pp. 34–36.
 10. Mukaeva L.A. Retrospektivnyi analiz dinamiki chislennosti naseleniya Chechenskoj Respubliki [A retrospective analysis of the population dynamics of the Chechen Republic]. *Vestnik Chechenskogo gosudarstvennogo universiteta*. 2015. no. 1, pp. 183–186.
 11. Polyakov E.M. Demograficheskaya situatsiya v sovremennoj Rossii [The demographic situation in modern Russia]. *Obshchestvo: sotsiologiya, psikhologiya, pedagogika*. 2012. No. 2, pp. 17–21.
 12. Prihod'ko S.B. *Doveritel'nyi interval nelineinoy regressii vremeni voss-tanovleniya rabotosposobnosti ustroystv terminal'noi seti* [Confidence interval of a nonlinear regression of time to restore the terminal network devices]. Khar'kov. 2014. no. 3/4(69), pp. 26–29.
 13. Prihod'ko S.B. *Ynterval'ne ocinjувanyja statystychnyh momentiv negausivs'kyh vypadkovykh velychyn na osnovi normalizujuchykh peretvoren'* [Intervalle ocenuvanje negassa statistical moments of random variables based on normalcy transformations]. *Matematychni modeljuvanyja*. Dni-prodzhzhyns'k: DDTU. 2011. No. 1(24), pp. 9–13.
 14. *Federal'naya sluzhba gosudarstvennoj statistiki* [Federal Service of State Statistics]. <http://www.gks.ru> (accessed 24.05.2017).
 15. Cherepanov E.V. *Veroyatnostno-statisticheskie osnovy prikladnoi sotsiologii i marketingovykh issledovanii* [Probabilistic and statistical foundations of applied sociology and marketing research]. Akad. Menedzhmenta Innovatsii, Kaf. “Analiticheskii menedzhment”, Byuro ekonomicheskikh i sotsial'nykh tekhnologii-BEST AMI. Moscow, 2006.
 16. Yakubovskii-Lipskii Yu.O. *Izmenenie komponentov obshchei chislennosti naseleniya na territorii rossijskoi federatsii s 1745 goda po 2010*

god [Changes in the components of the total population in the territory of the Russian Federation from 1745 to 2010]. *Ekologicheskaya ekspertiza*. 2013. no. 2, pp. 3–26.

ДАнные ОБ АВТОРАХ

Кочегарова Ольга Сергеевна, доцент кафедры «Математика и математическое моделирование», кандидат педагогических наук
ФГБОУ ВО Саратовский государственный аграрный университет имени Н.И. Вавилова

Театральная пл., 1, г. Саратов, 410012, Российская Федерация
kos19051979@gmail.com

SPIN-код: 3872-5780

Лажаннинкас Юлия Владимировна, доцент кафедры «Экономическая кибернетика», кандидат педагогических наук
ФГБОУ ВО Саратовский государственный аграрный университет имени Н.И. Вавилова

Театральная пл., 1, г. Саратов, 410012, Российская Федерация
lazhauninkas@yandex.ru

SPIN-код: 8735-6760

DATA ABOUT THE AUTHORS

Kochegarova Ol'ga Sergeevna, Associate Professor «Mathematics and Mathematical Modeling», Candidate of Pedagogical Sciences
Saratov State Agrarian University

1, Teatral'naya Sq., Saratov, 410012, Russian Federation
kos19051979@gmail.com

SPIN-code: 3872-5780

Lazhauninkas Julia Vladimirovna, Associate Professor «Economic Cybernetics», Candidate of Pedagogical Sciences
Saratov State Agrarian University

1, Teatral'naya Sq., Saratov, 410012, Russian Federation
lazhauninkas@yandex.ru

SPIN-code: 8735-6760