

Применение arima-моделей для краткосрочного прогнозирования показателей рождаемости в г. Уфа

Applying ARIMA-Models for Short-term Forecasting Fertility in the City of Ufa

Бахитова Р.Х., Лакман И.А., Шамсутдинова Н.К.

Bahitova, RH, Lakman, IA, Shamsutdinova, NK

Получено: 13.09.2016 г. Одобрено: 28.09.2016 г. Опубликовано: УДК 314 DOI: 10.12737/22310

Бахитова Р.Х.

д.э.н., заведующая кафедрой математических методов в экономике Башкирского государственного университета. Email: bakhitovarih@mail.ru

Лакман И.А.

к.т.н., доцент кафедры вычислительной математики и кибернетики Уфимского государственного авиационного технического университета. Email: Lackmania@mail.ru

Шамсутдинова Н.К.

к.и.н., с.н.с. Центра изучения человеческого потенциала ГАНУ «Институт стратегических исследований Республики Башкортостан». Email: shamsutdinovank@gmail.com

Bakhitova, RH, Lakman, IA, Shamsutdinova, NK

Bakhitova, RH

PhD in Economics, Head of the Chair of Math Methods in Economics, Bashkir State University. Address: 32 Zaki Validi Street, Ufa 450074, Republic of Bashkortostan, Russia. Email: bakhitovarih@mail.ru

Lakman, IA

PhD in Engineering Sciences, Associate Professor, Ufa State Aviation Technical University. Address: 12 Karl Marx Street, Ufa 450074, Republic of Bashkortostan, Russia. Email: Lackmania@mail.ru

Shamsutdinova, NK

PhD in History, Senior Research Worker, Centre of Human Potential Studies, Institute of Strategic Studies. Address: 20/1 50 Let Oktyabrya Street, Ufa 450076, Republic of Bashkortostan, Russia. Email: shamsutdinovank@gmail.com

Аннотация

Цель исследования - построение ARIMA-моделей относительных показателей рождаемости для г. Уфа – возрастных коэффициентов рождаемости и суммарного коэффициента рождаемости на период до 2021 г.

Предмет исследования – динамика относительных показателей рождаемости в г. Уфа. Результаты моделирования указывают на продолжение тенденции «взросления» материнства в г. Уфа. При сохранении тенденции роста уровня рождаемости в возрастных группах 25-44 года суммарный коэффициент будет расти. Согласно модели с расчетом доверительных интервалов, суммарный коэффициент может снизиться до уровня 2000-2001 гг.

Полученные результаты прогноза возрастных коэффициентов рождаемости могут быть использованы при построении сценариев рождаемости и дальнейшем расчёте численности и возрастно-половой структуры населения Уфы классическим когортно-компонентным методом.

Ключевые слова: рождаемость, прогноз, временные ряды, ARIMA-модель, г. Уфа.

Abstract

The Purpose of the Study is creating the ARIMA models of relative indicators of the birth rate for Ufa – age coefficients of the birth rate and total coefficient of the birth rate for the period up to 2021.

The Object of the Study is the dynamics of relative indicators of the birth rate in Ufa. The results of modeling indicate the continuation of a tendency of "growing" of motherhood in Ufa. When preserving a tendency of the growth of level in the birth rate for 25-44 the total coefficient will grow in age groups. According to the model with calculation of confidential intervals, the total coefficient can decrease. The received results of the forecast of age coefficients of the birth rate can be used for the creation of

scenarios of the birth rate and the further calculation of number and the age and sex structure of the population of Ufa by a classical cohort-component method.

Keywords: the birth rate, a forecast, time series, an ARIMA model

1. Введение

На первом этапе построения любых стратегий экономического развития города всегда должно анализироваться его демографическое развитие, так как стратегия должна ориентироваться на то, какая ожидается численность населения города в будущем, и какая будет половозрастная структура населения. Существует множество методов прогнозирования демографических показателей, центральное место в настоящее время принадлежит математическим методам, в частности, моделям временных рядов. Сгенерированные на их основе сценарии демографических процессов в дальнейшем используются для построения прогноза когортно-компонентным методом [7]. Однако в России такие методы до последнего времени были не столь распространены [2]. Один из таких методов - ARIMA-моделирование, который представлен в рекомендациях ООН для прогнозирования численности населения [1]. Он успешно используется в моделировании различных демографических процессов [11]. В то же время в России на региональном и муниципальном уровнях данная методология практически не используется. Следует оговориться, что ARIMA дает качественные прогнозы на относительно короткий период [8]. Поэтому наиболее вероятным прогнозным горизонтом полагаем период до 2021 года.

Целью данного исследования является построение прогноза по временным рядам возрастных коэффициентов рождаемости и суммарного коэффициента рождаемости. Данная научно-исследовательская работа велась в рамках реализации постановления Администрации городского округа города Уфа «О разработке Стратегии развития городского округа город Уфа Республики Башкортостан» для изучения перспективных параметров демографического развития, что повышает ее практическую значимость.

2. Методология прогнозирования

Первым этапом построения ARIMA-моделей является определение типа процесса, к которому относится временной ряд, на ретроспективе которого строится прогноз. Тип процесса (DS - процесс единичного корня ($x_t = x_{t-1} + \mu + \varepsilon_t$) или TS - процесс, содержащий детерминированный линейный тренд ($x_t = a + bt + \varepsilon_t$)) определяется на основе применения пятиэтапной процедуры расширенного теста Дикки-Фуллера [9]. Также определялся порядок интегрированности случайного процесса, описывающего временной ряд. Выяснилось, что процессы, описывающие ряды ретроспективных данных по возрастным коэффициентам рождаемости среди когорт 15-19, 25-29, 30-34 и 45-49 относятся к DS процессам первого порядка интеграции (I(1)). Для описания рядов динамики возрастных коэффициентов среди когорт 35-39 и 40-44 используется процесс I(2). Для возрастной когорты 20-24 применен процесс I(0).

Модели ARIMA (модели авторегрессии проинтегрированного скользящего среднего) имеет три порядка: p , d , q , где порядок p - порядок авторегрессии, порядок q - порядок процесса скользящей средней, порядок d отвечает за степень интеграции процесса. И соответственно вторым этапом ARIMA-моделирования является проведение процедуры идентификации или определения порядков моделей. Порядки модели ARIMA определялись согласно стандартной процедуры идентификации на основе анализа коррелограмм автокорреляционной и частной автокорреляционной функций [4], а также с помощью теста Лjung-Бокса при уровне значимости $\alpha=0,05$. Здесь p - порядок авторегрессии (максимальный лаг запаздывания в AR-процессе), q - порядок процесса скользящего среднего (максимальный лаг запаздывания в MA-процессе), d - порядок взятия разностей, то есть порядок интегрирования для DS-процесса.

Третьим этапом построения модели является параметризация или непосредственная оценка методом наименьших квадратов коэффициентов модели. К четвертому этапу можно отнести проведение процедуры селекции, основу которой составляли расчеты информационных критериев

Акайке, Шварца, Ханена-Квина [10] и выбиралась модель с минимальными значениями критериев.

Этап проверки качества построенной модели сводился к проведению процедуры проверки адекватности построенных прогнозных моделей, которая сама в свою очередь состояла из проведения подэтапов:

Расчет коэффициентов качества подгонки модели под реальные данные: коэффициент детерминации и скорректированный коэффициент детерминации. Коэффициенты отвечают за долю дисперсии, объясненной уравнением модели. Считается, чем ближе эта доля к единице, тем лучше качество оценки модели. Скорректированный коэффициент детерминации в отличие от коэффициента детерминации учитывает степени свободы в модели.

Проверка гипотезы о статистической значимости модели в целом, основанной на расчете F-критерия. F-статистика отвечает за тест, при нулевой гипотезе о статистической незначимости прогнозной модели в целом.

Проверка эффективности полученных методом наименьших квадратов сводится к проверке отсутствия автокорреляции в остатках модели на основе критерия Дарбина-Уотсона. Наличие автокорреляции в остатках приводит к неэффективным оценкам коэффициентов модели, и как следствие препятствует получению надежных результатов прогнозирования. Близость расчетного значения статистики Дарбина-Уотсона к 2 говорит об отсутствии автокорреляции, и, напротив, близость к 0 или 4 свидетельствует о наличии отрицательной или положительной автокорреляции соответственно.

Определение средней ошибки аппроксимации прогноза в процентах, которая служит для определения качества прогнозной модели.

Определение значения коэффициента неравенства Тейла. Близость значения коэффициента неравенства Тейла к нулю свидетельствует о хорошем качестве подгонки модели под исходные данные.

После проведения всех этапов построения модели, получают достоверные, с определенной долей допущения, результаты.

3. Прогнозирование возрастных коэффициентов рождаемости в Уфе

При построении прогноза возрастных коэффициентов рождаемости следует обязательно учитывать специфику демографического перехода. Происходят существенные изменения в жизненных стратегиях и репродуктивном поведении, молодые женщины в первую очередь стремятся получить образование и построить карьеру. Это, в свою очередь, способствует откладыванию создания семей и рождения детей, «взрослению» материнства. [3]. Для Республики Башкортостан и, в частности, для ее столицы, города Уфа, присущи демографические тенденции, наблюдаемые в России в целом.

Если анализировать ретроспективные данные по возрастным коэффициентам в городе Уфа в период за последние 25 лет (1990-2014 гг.), то можно отметить следующие тренды. Тенденцию снижения числа рождений возрастной когорте от 15 до 19 лет, в первую очередь, можно объяснить более широким распространением методов контрацепции. Снизилась рождаемость в когорте 20-24 года и выросла в группах 25-29 и 30-34 лет. В 2001-2003 гг. относительное число рождений в группах 20-24 и 25-29 лет сравнялось, затем стала опережать более старшая группа. В результате возраст максимальной репродуктивной активности сместился с группы 20-24 года в группу 25-29 лет. В масштабах России это случилось в 2008 г. [6]. К 2010 г. обошла группу 20-24-летних и более возрастная когорта - 30-34 летних. И в возрастных когортах 35-39 и 40-44 лет по сравнению с 1990-ми годами, произошел рост. Минимальный вклад в рождения вносит группа старше 45 лет.

Рост интенсивности рождений в последнее десятилетие определенно связан с улучшением социально-экономической ситуации и введением мер поощрения рождаемости в 2007 г. [6]. Повышение рождаемости в старших группах связано, скорее всего, и с реализацией отложенных в 1990-е гг. рождений, и распространением более позднего материнства в новом веке.

Результаты применения ARIMA моделирования показали удовлетворительные статистические результаты, что повысило надежность полученных прогнозных оценок динамики коэффициентов рождаемости у женщин репродуктивного возраста в пятилетних когортах.

Согласно построенным прогнозам, в возрастных группах 15-19 и 45-49 лет. интенсивность рождаемости практически не будет меняться, их вклад в рождения будет минимальным. Ожидается дальнейшее снижение рождаемости в когорте 20-24 -летних, и повышение в более старших возрастах: 25-29, 30-34, 35-39, 40-44 лет - на 22,76; 42,8; 51,7; 32,8% соответственно.

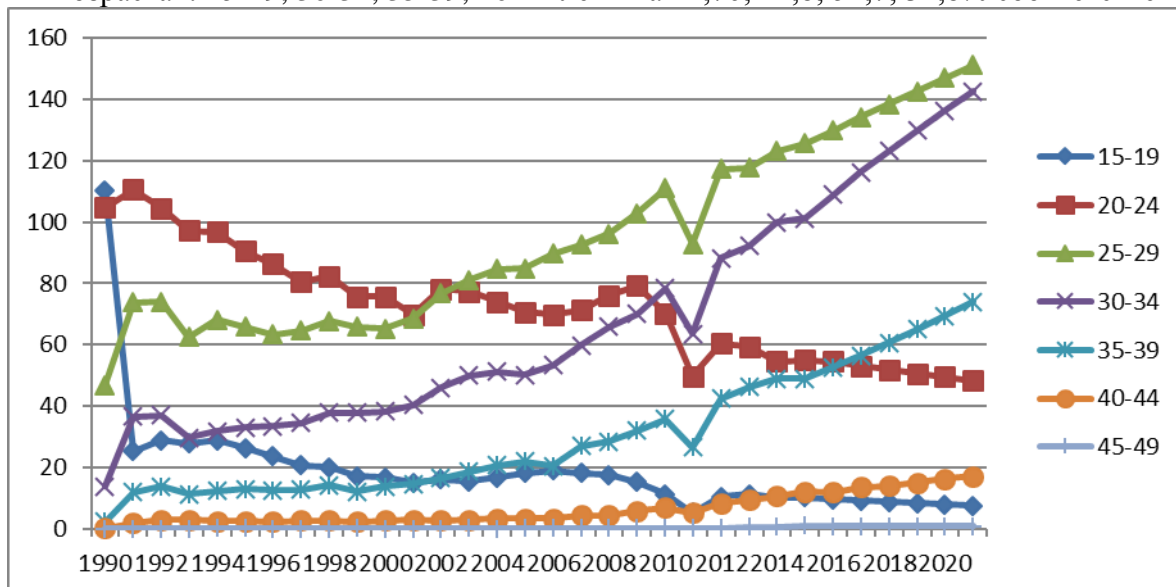


Рис. 1. Возрастные коэффициенты рождаемости в г.Уфа: и фактические (1990-2014 гг.) и прогнозируемые (2015-2021 гг.)

Рассмотренные тенденции изменений возрастных коэффициентов рождаемости способствовали росту значения суммарного коэффициента рождаемости, в том числе в г.Уфа, который, при их продолжении, обусловит дальнейшее его повышение (рис.2).

Достигнувший "дна" в 2000-2011 гг., суммарный коэффициент в г.Уфа в имел устойчивую тенденцию повышения в 2000-е гг.. Расчет его значений на основе полученных возрастных коэффициентов рождаемости указывает на возможность его дальнейшего повышения (рис. 2).

Возможно, данный рост в старших возрастах будет менее значительным, поскольку, как уже было сказано, свой вклад в повышение в 2000-е гг. внесла реализация отложенных рождений. В 2007-2013 гг. росла вероятность вторых и третьих рождений, но первых – снижалась, то есть база для реализации вторых и третьих рождений сужалась [6].. В свою очередь, это говорит о существовании ограничений дальнейшего повышения возрастных коэффициентов рождаемости в старших возрастах.

Моделирование динамики суммарного коэффициента рождаемости, проведенное с расчетом доверительных границ в два стандартных отклонения, показывает менее оптимистичные сценарии за счет большой ширины доверительного интервала.

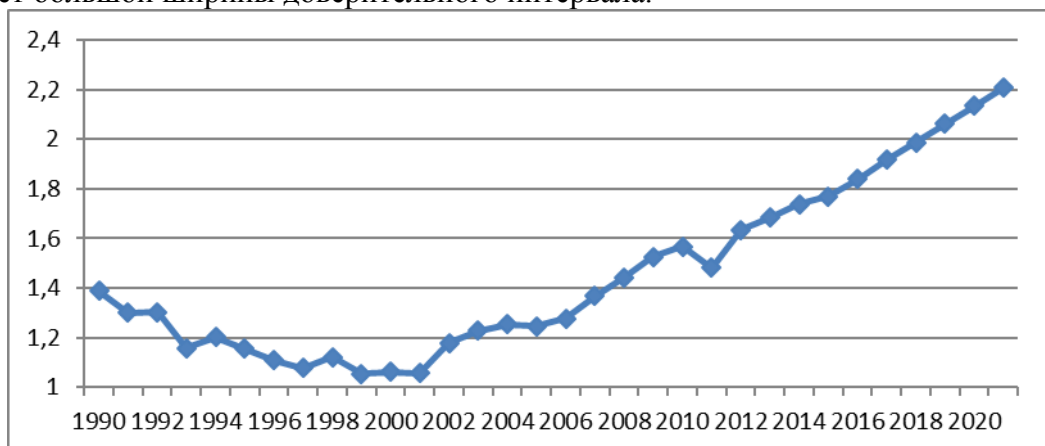


Рис. 2. Динамика суммарного коэффициента рождаемости в г.Уфа фактические (1990-2014 гг.) и прогнозируемые (2025-2021 гг.).

Нижняя граница доверительного интервала указывает на возможность снижения значения суммарного коэффициента рождаемости до уровня 2001 г., верхняя граница - на дальнейшее повышение. Средние показатели не слишком отличаются от сегодняшнего уровня.

4. Выводы

Таким образом, моделирование динамики возрастных коэффициентов рождаемости в г.Уфа указывает на сохранение тенденции "взросления" материнства и роста рождаемости в когортах старше 25 лет. Результаты указывают на рост суммарного коэффициента. Средний точечный прогноз указывает на возможность "стагнации" или даже снижения суммарного коэффициента рождаемости.

Использование математических методов в построении сценариев будущих изменений рождаемости остается актуальной научно-практической задачей. Практическая значимость исследования заключается и в том, что полученные результаты моделирования возрастных коэффициентов рождаемости и суммарного коэффициента рождаемости могут быть использованы при построении сценариев рождаемости и дальнейшем расчете численности и возрастно-половой структуры населения Уфы классическим когортно-компонентным методом.

Литература

1. Голдстоун Дж., Шульги С.Г., Коротаев А.В., Архангельский В.Н., Зинькина Ю.В, Новиков К.Е., Пустовалов Д.Н. Политическая демография России. Политика и государственное управление. – М.: Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации, 2015.
2. Захаров С.В. Обобщенная модель воспроизводства населения и ее значение для практики демографического анализа// Экономический журнал ВШЭ. – 2011. – № 4. – С. 409-443.
3. Захаров С.В., Фрейка Т. Эволюция рождаемости в России за полвека: оптика условных и реальных поколений// Демографическое обозрение. – 2014. – №1. Т. 1. – С. 106-143.
4. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс.- М.: Дело, 2007. – 504 с.
5. Население России 2009. Семнадцатый ежегодный демографический доклад / отв. ред. А. Г. Вишневский; Нац. исслед. ун-т «Высшая школа экономики». – М. : Изд. дом Высшей школы экономики, 2011. – 334 с.
6. Население России 2013. Двадцать первый ежегодный демографический доклад. Ответственный редактор С.В. Захаров. – М.: Изд. дом Высшей школы экономики, 2015. – 428 с.
7. Энциклопедия статистических терминов. В 8 т. Т. 5. Демографическая и социальная статистика. – М.: Росстат, 2011.
8. De Beer J. Projecting age-specific fertility rates by using time-series methods// Eur J Popul, 1990. Mar. 5(4). -P. 315-46.
9. Dickey D. A. and Fuller W. A. Distribution of the Estimators for Autoregressive Time Series with a Unit Root / Journal of the American Statistical Association. 74. 1979. - p. 427- 431.
10. McQuarrie A. D. R., Tsai C. L. Regression and time series model selection.– World Scientific, 1998. – 455 с.
11. Readings in Population Research Methodology/ Project editors: Bogue D.J., Arriaga E.E., Anderton D.L. Vol. 5. Population Models, Projections and Estimates. Chicago (Illinois): United Nations Fund for Population Activities, 1993.
12. Stenley K Smith, Jeff Tayman, David A Swanson. State and Local Population Projections;: Methodology and Analysis/ Kluwer Academic Publishers: New York, Boston, Dardrecht, London, Moscow, 2002. - 443 p

References

1. Goldstoun Dzh., Shul'gi S.G., Korotaev A.V., Arhangel'skij V.N., Zin'kina Ju.V, Novikov K.E., Pustovalov D.N. Politicheskaja demografija Rossii. Politika i gosudarstvennoe upravlenie. – М.:

- Rossijskaja akademija narodnogo hozjajstva i gosudarstvennoj sluzhby pri Prezidente Rossijskoj Federacii, 2015.
2. Zaharov S.V. Obobshhennaja model' vosproizvodstva naselenija i ee znachenie dlja praktiki demograficheskogo analiza// Jekonomicheskij zhurnal VShJe. – 2011. – № 4. – S. 409-443.
 3. Zaharov S.V., Frejka T. Jevoljucija rozhdaemosti v Rossii za polveka: optika uslovyh i real'nyh pokolenij// Demograficheskoe obozrenie. – 2014. – №1. T. 1. – S. 106-143.
 4. Magnus YA.R., Katyshev P.K., Pereseckij A.A. EHkonometrika. Nachal'nyj kurs.- M.: Delo, 2007. – 504 s.
 5. Naselenie Rossii 2009. Semnadcatyj ezhegodnyj demograficheskij doklad / otv. red. A. G. Vishnevskij ; Nac. issled. un-t «Vysshaja shkola jekonomiki». – M. : Izd. dom Vysšej shkoly jekonomiki, 2011. – 334 s.
 6. Naselenie Rossii 2013. Dvadcat' pervyj ezhegodnyj demograficheskij doklad. Otvetstvennyj redaktor S.V. Zaharov. – M.: Izd. dom Vysšej shkoly jekonomiki, 2015. – 428 s.
 7. Jenciklopedija statisticheskikh terminov. V 8.t. T. 5. Demograficheskaja i social'naja statistika. – M.: Rosstat, 2011.
 8. De Beer J. Projecting age-specific fertility rates by using time-series methods// Eur J Popul, 1990. – Mar. 5(4). – P. 315-46.
 9. Dickey D. A. and Fuller W. A. Distribution of the Estimators for Autoregressive Time Series with a Unit Root / Journal of the American Statistical Association. 74. 1979. - P. 427- 431.
 10. McQuarrie A. D. R., Tsai C. L. Regression and time series model selection. – World Scientific, 1998. – 455 p.
 11. Readings in Population Research Methodology/ Project editors: Bogue D.J., Arriaga E.E., Anderton D.L. Vol. 5. Population Models, Projections and Estimates. (Illinois): United Nations Fund for Population Activities, 1993.
 12. Stenley K Smith, Jeff Tayman, David A Swanson. State and Local Population Projections:; Methodology and Analysis/ Kluwer Academic Publishers: New York, Boston, Dardrecht, London, Moscow. 2002. – 443 p.