The Lee–Carter Method for Mortality Forecasting: the Case of the Republic of Bashkortostan*

Irina Lakman[1 [0000−0001−9876−9202]], Denis Popov[2 [0000−0002−7698−8858]], and Nailya Shamsutdinova[2 [0000−0003−2008−0856]]

[1] Ufa State Aviation Technical University, Ufa, Russia,
lackmania@mail.ru,
[2] Institute for Strategic Studies of the Republic of Bashkortostan, Ufa, Russia

Abstract. The article is devoted to predicting the mortality rates by age and sex for one of constituent entities of the Russian Federation – the Republic of Bashkortostan. As initial data, age-specific rates were used in five-year groups of up to 85 years, published by the Territorial Body of the Federal State Statistics Service for the Republic of Bashkortostan (2001–2014) and calculated by the authors (1980–2000). The death rate was calculated by means of Stata software. For this purpose, the method was specifically adapted for the Republic of Bashkortostan. However, correlation of the actual indices of 2015 showed that predicted values for that year were underestimated. Due to the impact of downward dynamic observed nationwide in the 1990s, there are restrictions on using of methods of extrapolation of death rate for the purpose of population forecasts and calculation of insurance risks and decision-making in this field. The outcome for the Republic of Bashkortostan, as for Russia in whole, indicates that there still remains a crucial task of reducing economic losses and losses of human capital as a result of high mortality of the working-age population, which will bring pressure upon pension funds.

Keywords: Insurance risks, mortality forecasting rate, death rate, the Lee–Carter method, ARIMA model

1 Introduction

In the Russian Federation, after a period of high rise of mortality rate during 1990s, the situation changes into upward dynamics since 2004. From this time onward until 2016, life expectancy at birth increases in figures nationwide totaled up to 6.6 years.

The dynamic pattern of life expectancy at birth on a global scale has the tendency to linear growth; however, in Russia this trend has extremely unstable

* The reported study was funded by Russian Foundation for Basic Research (RFBR) according to the research project No 17-46-020237 p_a.

character [15]. The increase in life expectancy is one of the main goals of the socioeconomic and demographic policy of the Russian Federation and its regions. The Republic of Bashkortostan as one of the large and influential constituent entity of the Russian Federation, replicates the whole Russia's situation on the dynamics of mortality rate. The study of the applicability of prediction methods at the regional level is now important not only from scientific and practical viewpoints.

Age-specific mortality rate is the basis for predicting life expectancy, the future population, its structure, and also of interest to the insurance industry. Within the framework of human capital research, studying these particular age groups, where the greatest loss rates are being observed, over time it will be possible to create a system for loss prevention. Besides, the mortality forecasting is necessary for model building growth prospects of the insurance and pension systems, contributing decision-making in the range of acceptable risk.

2 Overview of actual methodological and instrumental approaches

In the estimation of studies of Sweden, the USA and some other countries, pension funds are subject to consider longevity risk, which actualizes the search for new methods for their management [14, 16].

At the other end of the scale, premature mortality decreases the efficiency of investments into human capital and serves as the reason of economic damage that is currently important for Russia [10]. According to certain estimates, it amounted to 16.3% of Gross Regional Product (GRP) in the Republic of Bashkortostan [3]. In order to assess the future prospects and choose priorities for mortality reduction, it is necessary to resort to modeling its indicators.

For forecasting age-specific mortality index, the Lee–Carter method is widely used [5,8,9,12,13]. The results attained across the Russian Federation show continued existence in the future of a big difference between male and female mortality nationwide in comparison with other countries. However, high fluctuations of mortality in Russia demand the questioning attitude towards the results of application the Lee–Carter method [10, 11]. However, in the regions of the Russian Federation, no simulation experiment was performed using the Lee–Carter model. Coherent forecasts of mortality industrialized countries are justified by reason of their greater homogeneity [4]. The Russian Federation continues to be extremely internally heterogeneous [17].

The Lee–Carter method original modeling contains the following equation:

$$\log m_{x,t} = a_x + b_x k_t + \varepsilon_{x,t}$$

Here $m_{x,t}$ – age-specific mortality index for a cohort x in the period of time t (year), a vector reflects time-mean value effect of influence of age to mortality index for each cohort x, k_t vector reflects an effect of influence of time, average on age, of mortality index for every period of time t, and the coefficient vector

b_x explains the effect of interaction expressing specific sensitivity of mortality index at age x to changes in time of k [16]. Random errors of the equation are designated as $\varepsilon_{x,t}$.

Coefficient vectors b_x and k_t are in the original model of the Lee–Carter method by means of singular value decomposition (SVD) which equates:

$$A = USV^T$$

Here U – the orthogonal matrix, V^T – the transposed matrix,S – the diagonal matrix consisting of zero and singular values of matrix A spaced diagonally. There is a use of experience for finding unknown b_x and k_t as an alternative to SVD-analysis of the weighted least spreads method (WLS) [7] and method of maximum likelihood (ML) [6].

The first base of index construction of ARIMA models (autoregressive integrated moving average) is type definition of process to which time-series belongs. The approach of J. J. Dolado , T. Jenkinson and S. Sosvilla-Rivero, consistently applying to the complex hypotheses of the extended test of Dickey–Fuller allows to determine the type of process. Further, if there is a deterministic linear component in the time series, it is removed. If the time series is an integrable process in the first or second order, then the procedure for differentiation of the corresponding order is performed.

At the second stage, the identification procedure of an order of autoregression and order of process of the moving average is carried out.

At the third stage, the ARIMA model equation coefficients are estimated by method of least squares and calibrate reversibility of model.

At the third stage, the coefficients of ARIMA model are estimated by the least squares method and they verify the reversibility of the model, which means, they test the requirement that the roots of the characteristic process corresponding to the process lie outside the unit circle.

At the fourth stage, selection of models by using information criteria of Akaike, Hannah–Quinn is carried out in case the same process can be described in various equations.

At the fifth stage, quality monitoring of the constructed model is conducted, screening with the help of specific tests (Jarque–Bera, Durbin–Watson, Breusch–Godfrey, etc.) so that estimated coefficients of ARIMA model could be unbiased, well-founded and effective.

At the final stage, forecasting model behavior is estimated, proceeding from a minimum mean absolute percent error, residual dispersion and odds ratio according to Theil inequality (index).

3 Materials and alternatives

For the development of population mortality forecasting by sex-age structure, customized 5-year age groups of up to 85+ cohorts study has been set up. From 2001 to 2014 sex-age-specific death rates appeared in official publications [1, 2].

From 1980 to 2000 the interval charts were provided by Territorial Body of Federal State Statistics Service in the Republic of Bashkortostan (Bashkortostanstat, Ufa, Russia), calculated by a standard method:

$$m_x = \frac{M_x}{\bar{P}_x} \times 1000 \tag{1}$$

Here m_x – age-specific mortality index; – number of the deceased aged x in a year; $\bar{P}_x$ – mid-year population aged x. Change of mortality age distribution in the Republic of Bashkortostan, as in Russia at large, has irregular nature. Tremendous losses in the 1990s were suffered by the population at productive age (employable age). Increase in remaining life expectancy in the country continues since 2003. However, in this particular region the death rate advances by 2015 in comparison with 1990 in some age groups: among women in the cohort of 25–39 years and among men in the cohort of 25–49 years and 60–74 years old (Tables 1 and 2).

Table 1. The age-specific male death rate from 1990 to 2015

Kohort	1990	1995	2000	2005	2010	2015
0–1	18.3	21.5	17.5	14.4	7.7	8.0
1–4	1.4	1.2	1.1	0.8	0.6	0.3
5–9	0.7	0.8	0.6	0.5	0.3	0.3
10–14	0.5	0.8	0.5	0.5	0.4	0.3
15–19	1.5	2.3	2.5	1.6	1.3	1.0
20–24	2.8	5.0	5.5	3.9	3.2	2.4
25–29	3.2	6.0	6.0	5.8	5.8	3.5
30–34	3.8	7.1	7.3	7.4	7.7	6.4
35–39	4.9	9.3	8.5	9.3	8.2	9.1
40–44	6.8	11.4	11.1	12.9	9.7	10.0
45–49	10.0	15.6	14.7	16.0	13.0	13.1
50–54	13.9	21.6	19.8	22.2	17.5	16.6
55–59	20.1	26.4	28.3	29.9	23.8	23.0
60–64	28.9	37.1	38.4	42.1	35.9	33.4
65–69	41.0	51.7	52.9	53.2	49.1	43.1
70–74	60.5	69.9	71.3	74.9	66.8	62.0
75–79	90.3	101.5	94.9	104.2	97.0	87.2
80–84	135.0	149.4	132.1	142.8	136.6	125.0
85+	213.8	248.5	238.2	235.3	204.3	193.8

4 Experimental setup

After receiving coefficients b_x and k_t by means of singular value decomposition executed using Stata software, parameterization of ARIMA models for k_t series

Table 2. The age-specific female death rate from 1990 to 2015

Kohort	1990	1995	2000	2005	2010	2015
0–1	13.9	15.1	12.0	9.4	6.6	6.8
1–4	1.0	1.0	0.8	0.7	0.5	0.3
5–9	0.4	0.4	0.4	0.3	0.2	0.2
10–14	0.3	0.4	0.3	0.2	0.3	0.3
15–19	0.8	0.9	0.7	0.8	0.7	0.5
20–24	1.0	1.0	1.1	0.9	0.8	0.8
25–29	0.8	1.3	1.3	1.4	1.5	1.3
30–34	1.0	1.4	1.5	1.9	2.1	2.2
35–39	1.4	2.1	2.2	2.6	2.3	2.9
40–44	2.1	3.4	2.9	3.5	3.2	3.2
45–49	3.8	5.1	4.2	4.8	4.2	4.2
50–54	5.3	7.3	6.9	7.0	5.3	5.7
55–59	8.2	9.8	9.8	11.0	9.2	7.4
60–64	12.4	15.7	14.7	15.8	13.4	11.1
65–69	19.4	23.5	23.0	22.7	20.8	17.0
70–74	32.8	35.7	37.3	38.5	32.0	27.9
75–79	53.1	63.3	59.2	62.6	56.2	47.6
80–84	88.4	105.5	101.6	100.7	92.9	84.2
85+	181.5	203.6	198.6	214.6	184.2	171.4

by sex distribution was carried out. In Fig. 1 and 2, k_t time series are shown calculated for men and women respectively.

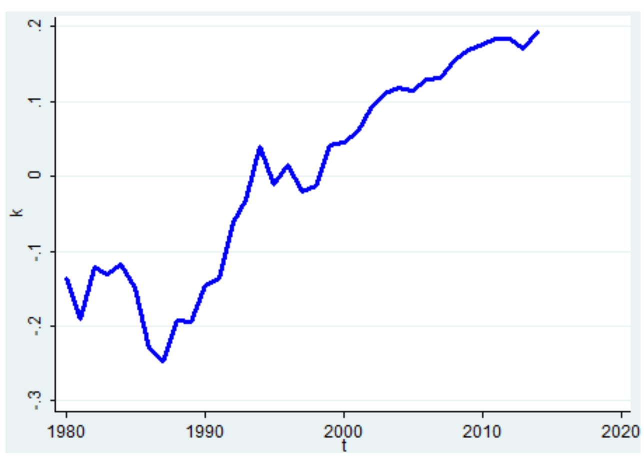

Fig. 1. Male death rate (k_t)

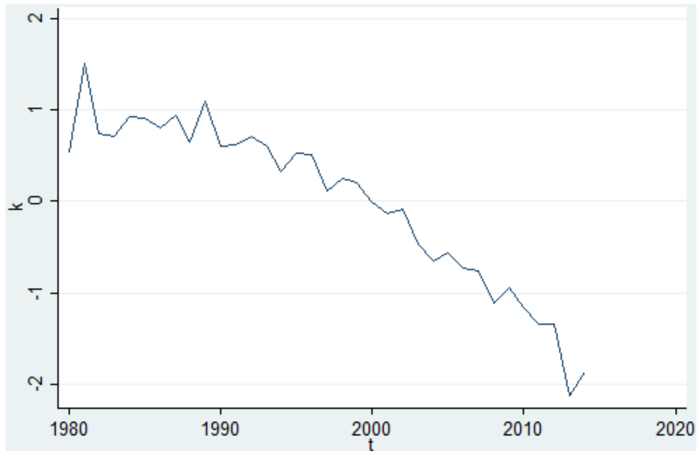

Fig. 2. Female death rate (k_t)

As a result, the best model for estimation of k_t on male death rate is ARIMA (2,1,2) model:

$$\Delta k_t = 0.023 \cdot \Delta k_{t-1} - 0.724 \Delta k_{t-2} - 0.214 \cdot \varepsilon_{t-1} - 0.04 \cdot \varepsilon_{t-2} + 0.01 + \varepsilon_t$$

The best model for estimation of k_t on female death rate is the ARIMA (1,2,2) model:

$$\Delta^2 k_t = -0.381 \cdot \Delta^2 k_{t-1} - 0.862 \cdot \varepsilon_{t-1} - 0.871 \cdot \varepsilon_{t-2} - 0.005 + \varepsilon_t$$

Both models were checked for lack of residual autocorrelation by Ljung–Box Q-test and for normality of their distribution by Jarque–Bera test.

The Lee–Carter models designed for other countries have k_t process as random-walk process in terms of DS (I (1)) with a constant [6]. For the Republic of Bashkortostan k_t process serves as a framework for predictive model of the Lee–Carter method, presents the actual process of DS (I (2)) with a constant and AR component (AR (1)) for women and DS (I (1)) with a constant and AR component (AR (2)) for men. This indicates clearly the difference of mortality dynamics nature in the regions of Russia and in other countries.

Within the scope of ARIMA (2,1,2) and (1,2,2) models, forecasts for k_t for the period from 2015 to 2030 has been generated.

On the basis of the ARIMA (2,1,2) and (1,2,2) models, forecasts for k_t for the period from 2015 to 2030 were constructed (Tables 3 and 4).

5 Conclusion

The simulation observations point out persistent growth of mortality rate for men in all five-year cohorts from 25 to 74 years and for women from 25 to 49 years. Noticeable reduction index follows in both genders in group of 85+ years.

Table 3. Forecasted death rates for male population from 2015 to 2025

Kohort	2015	2016	2019	2022	2025
0–1	9.7	9.5	8.8	8.2	7.6
1–4	0.3	0.3	0.2	0.2	0.1
5–9	0.3	0.3	0.3	0.3	0.3
10–14	0.4	0.4	0.4	0.4	0.4
15–19	1.7	1.7	1.7	1.8	1.8
20–24	4.0	4.0	4.1	4.2	4.3
25–29	6.1	6.2	6.4	6.7	6.9
30–34	8.6	8.8	9.3	9.8	10.4
35–39	10.0	10.1	10.6	11.1	11.6
40–44	12.4	12.6	13.0	13.5	14.1
45–49	15.7	15.9	16.4	16.9	17.4
50–54	21.4	21.6	22.2	22.8	23.4
55–59	28.7	29.0	29.7	30.5	31.2
60–64	40.5	40.8	41.8	42.9	43.9
65–69	52.5	52.8	53.6	54.5	55.4
70–74	71.5	71.7	72.4	73.1	73.8
75–79	100.3	100.5	101.0	101.5	102.0
80–84	137.7	137.6	137.2	136.8	136.4
85+	206.7	204.8	199.6	194.4	189.4

Table 4. Forecasted death rates for female population from 2015 to 2025

Kohort	2015	2016	2019	2022	2025
0–1	5.3	4.9	3.9	3.1	2.4
1–4	0.6	0.6	0.5	0.5	0.5
5–9	0.2	0.1	0.1	0.1	0.1
10–14	0.2	0.2	0.1	0.1	0.1
15–19	0.6	0.6	0.5	0.5	0.5
20–24	0.9	0.9	0.9	0.9	0.9
25–29	1.6	1.6	1.8	2.0	2.2
30–34	2.7	2.8	3.3	4.0	4.8
35–39	3.0	3.1	3.4	3.8	4.3
40–44	3.5	3.6	3.8	4.0	4.2
45–49	4.5	4.5	4.6	4.6	4.7
50–54	6.1	6.1	6.1	6.1	6.1
55–59	9.6	9.6	9.7	9.8	9.9
60–64	13.7	13.7	13.7	13.7	13.7
65–69	20.3	20.3	20.1	20.0	19.8
70–74	31.1	30.6	29.3	27.9	26.5
75–79	56.3	56.3	56.1	55.9	55.7
80–84	93.1	92.9	92.1	91.4	90.5
85+	179.9	178.5	174.3	169.8	165.1

In spite of the fact that positive changes were outlined in separate groups of active working-age in recent years, the Lee–Carter model lets us see increase in practically all cohorts. In case if trends of continued existence in the last 34 years remain, death rate only in children and advanced ages may decrease. The projected values of age-specific mortality rates received by means of the Lee–Carter model application indicate preserving of considerable inequality of cohorts. The use of the Lee–Carter model for mortality forecasting makes it possible to concentrate attention on specific problematic age groups, refraction of situation in which allows avoiding evolvement of the negative scenario.

The medium-term projected perspective indicates continuance of male supermortality problem. However, the circumstances with female mortality rate are not so positive either.

The results by applying the Lee–Carter model should be considered particularly in terms of Russia, its mortality factor is needed to be explained exceptionally. They have an effect of nonlinear dynamics of death rate due to dramatic discontinuity during compound crisis in the 1990s. It is necessary to mention that, in the first place, this break was observed among the population at employable age. The common trend of infant and child mortality, declining throughout 1990s made an impact on the results of modeling which indicated continuation of this trend. It is noteworthy that the cohort of 30–34 years happens to be the highest possible death rates for both sexes. In accordance with Russian research, the greatest contribution falls in the lost years of potential life in some regions of Russia [10]. We have the opportunity to correlate the received forecast results with the actual results for 2015. Predicted indices turned out to be less optimistic.

There exists dozens of ways to consider the non-linearity of the perspective dynamics of mortality, besides all of them contribute to reasonably accurate predictions of mortality and, respectively, the remaining life expectancy [13]. A lack of the Lee–Carter model has been noted in the form of constant-rate of reducing mortality, which leads to overestimation of the future level of mortality, especially in older age groups [4]. In our case, the Lee–Carter model pointed to an underestimation of reduction of mortality which is related to the inconsistent fluctuation of death rate. Due to the impact of downward dynamic observed nationwide in the 1990s, there are restrictions on using methods of extrapolation of death rate for the purpose of population forecasts and calculation of insurance risks and decision-making in this field. The outcome for the Republic of Bashkortostan, as for Russia in whole, indicates that there still remains a crucial task of reducing economic losses and losses of human capital as a result of high mortality of the working-age population, which will bring pressure upon pension funds.

References

1. Demographic processes in the Republic of Bashkortostan (2001–2014). Stat. sb. Ufa, Russia

2. Population mortality of the Republic of Bashkortostan (2009–2014). Bashkortostanstat. Stat. sb. Ufa, Russia
3. Rating of regions for economic losses from premature death. Raiting agency "RIA Rejting": Mediagroup "Rossiya segodnya", http://riarating.ru/regions_rankings/20141216/610640551.html
4. Bergeron-Boucher, M.P., Canudas-Romo, V., Oeppen, J., Vaupel, J.W.: Coherent forecasts of mortality with compositional data analysis. Demographic Research 37(17), 527–566 (2017)
5. Danesi, I., Haberman, S., Millossovich, P.: Forecasting mortality in subpopulations using lee–carter type models: A comparison. Insurance: Mathematics and Economics 62(Supplement C), 151 – 161 (2015)
6. Haberman, S., Russolillo, M.: Lee-carter mortality forecasting: application to the ital-ian population. Actuarial Research Paper (2005)
7. Koissi, M., Shapiro, A., Hognas, G.: Fitting and forecasting mortality rates for nordic countries using the lee–carter method. Actuarial Research Clearing House 4(21) (2005)
8. Lee, R.D., Carter, L.R.: Modeling and forecasting u.s. mortality. Journal of the American Statistical Association 87(419), 659–671 (1992)
9. Li, J.: An application of MCMC simulation in mortality projection for populations with limited data. Demographic Research 30(1), 1–48 (2014)
10. Samorodskaya, I., Vatolina, M., Bojcov, S.: Methodical topics and results of assessment of a global burden of diseases (review of literature). Profilakticheskaya meditsina 18(1), 40–45 (2015)
11. Scherbov, S., Ediev, D.: Does selection of mortality model make a difference in projecting population ageing? Demographic Research 34(2), 39–62 (2016)
12. Shang, H.L.: Point and interval forecasts of age-specific life expectancies: A model averaging approach. Demographic Research 27(21), 593–644 (2012)
13. Stoeldraijer, L., van Duin, C., van Wissen, L., Janssen, F.: Impact of different mortality forecasting methods and explicit assumptions on projected future life expectancy: The case of the Netherlands. Demographic Research 29(13), 323–354 (2013)
14. Tan, K.S., Blake, D.P., MacMinn, R.D.: Longevity risk and capital markets: The 2013–14 update. Pensions Institute Discussion Paper PI-1502 (2015)
15. Vishnevsky, A.: Mortality in Russia: the second epidemiologic revolution that never was. Dempographic Review 4 (2014)
16. Wong, J., Forster, J., Smith, P.: Bayesian mortality forecasting with overdispersion. University of Southampton Working Paper (2016)
17. Zubarevich, N.: Regional inequality and potential for modernization. In: Oxenstierna, S. (ed.) The Challenges for Russia's Politicized Economic System, pp. 182–201. Routledge, UK (2015)