

基于Keras的LSTM网络在死亡率预测的应用

肖鸿民, 李 祥

西北师范大学数学与统计学院, 甘肃 兰州

收稿日期: 2021年9月22日; 录用日期: 2021年10月15日; 发布日期: 2021年10月26日

摘 要

随着人口老龄化程度进一步加深, 加之人口死亡率的持续下降和人口预期寿命的不断延长, 对死亡率有更加有效的预测, 在长寿风险的研究中尤为重要。在以往的死亡率预测中, 死亡率模型的时间项大多采用传统的ARIMA时间序列方法进行预测。本文选取中国人口死亡率数据, 以长短期记忆(LSTM)网络 and 全连接层为基础, 构建基于循环神经网络的LSTM学习机, 用于预测Lee-Carter模型的时间项。

关键词

死亡率预测, 循环神经网络, 时间序列分析, LSTM

Application of LSTM Networks Based on Keras in Mortality Prediction

Hongmin Xiao, Xiang Li

College of Mathematics and Statistics, Northwest Normal University, Lanzhou Gansu

Received: Sep. 22nd, 2021; accepted: Oct. 15th, 2021; published: Oct. 26th, 2021

Abstract

As the aging degree of the population deepens, coupled with the continuous decline in population mortality and prolonged life expectancy, a more effective prediction of mortality is particularly important in the study of longevity risk. In previous mortality predictions, the time terms of mortality models were mostly predicted using traditional ARIMA time series methods. This paper selects Chinese population mortality data, based on long-short term memory (LSTM) network and dense layers, constructing a LSTM learning machine based on recurrent neural network to predict the time term of the Lee-Carter model.

Keywords

Mortality Prediction, Recurrent Neural Network, Analysis of Time Series, LSTM

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

国家统计局 2021 年 5 月 11 日公布第七次全国人口普查主要数据结果。数据显示,全国人口共 141,178 万人,与 2010 年(第六次全国人口普查数据)的 133,972 万人相比,增加了 7206 万人,增长率为 5.38%,平均增长率为 0.53%,比 2000 年到 2010 年的年平均增长率 0.57% 下降 0.04 个百分点。近 10 年来我国人口继续以低速态势保持增长。其中,0~14 岁人口为 25,338 万人,占比 17.95%;15~59 岁人口为 89,438 万人,占比 63.35%;60 岁及以上人口为 26,402 万人,占比 18.70% (其中,65 岁及以上人口为 19,064 万人,占比 13.50%)。与 2010 年相比,0~14 岁、15~59 岁、60 岁及以上人口的比重分别上升 1.35 个百分点、下降 6.79 个百分点、上升 5.44 个百分点。我国少儿人口比重回升,生育政策调整取得了积极成效。与此同时,更应该关注的应是人口老龄化程度进一步加深,加之人口死亡率的持续下降和人口预期寿命的不断延长,未来一段时期将持续面临人口长期均衡发展的压力。

死亡率作为决定人口规模、结构和分布变化的主要因素之一,对其预测显得尤为重要。死亡率预测应用在许多方面,并且有其特殊的意义。其中,在长寿风险领域,作为确定其成员年度缴款以及人口预测。由于其重要性,近几十年来,死亡率预测技术有了很大的发展和提高,涌现了多形式的死亡率模型。在国外, Lee 等(1992) [1]提出了一种将人口统计模型和时间序列联系在一起的对数双线性模型,即 Lee-Carter 模型(以下简称 LC 模型)。Renshaw 等(2006) [2]在 LC 模型的基础上加入了出生年效应,首次提出了 RH 模型并将其应用到英格兰和威尔士男性死亡率预测中。Cairns 等(2011) [3]通过对 6 个随机死亡率模型的分析,发现要从拟合和预测两个方面来判定一个模型的合理性。Li 等(2013) [4]运用扩展的 LC 模型方法对美国和日本死亡率数据进行拟合与预测。Enchev 等(2017) [5]利用 6 个国家的死亡率数据对共同年龄效应(CAE)等随机死亡率模型的拟合效果进行比较,通过模型预测发现 CAE 模型最优。Li 等(2019) [6]运用一个新的双人口死亡率预测模型对比利时、瑞典、瑞士和英国 4 个国家的高龄死亡率数据进行了拟合与预测。在国内研究中,李志生等(2010) [7]运用 2019 对中国人口死亡率数据进行拟合和预测,之后黄顺林等(2010) [8]运用 RH 模型对中国男性人口死亡率进行了预测,并将其应用到预期寿命和年金系数的估计中。韩猛等(2010) [9]对 LC 模型进行了改进,通过一个双随机过程对 LC 模型中的时间项进行建模。曾燕等(2016) [10]采用 Bootstrap 方法与 LC 模型拟合死亡率数据,运用带漂移项的随机游走模型进行预测。LC 模型对中国人口死亡率进行预测后,根据预测的死亡率数据对中国人口平均预期寿命进行计算。樊毅和张宁等(2018) [11]基于中国人口死亡率数据,对 8 个死亡率经典模型进行了比较分析,发现 APC 模型对中国人口死亡率的拟合效果最优。王晓军等(2019) [12]运用台湾高龄人口死亡率数据对贝叶斯分层模型等 8 种死亡率模型进行了比较分析,最后通过拟合中国大陆人口数据,发现 CBD 模型对我国大陆地区高龄死亡率的拟合和预测较好。

关于中国人口数据的死亡率模型研究,大多依赖于 LC 和 APC 等几个常见的随机死亡率模型。马海飞等(2020) [13]对 APC 模型进行扩展,提出一个新的死亡率模型 EAPC (Extended Age-Period-Cohort),通

过比较模型的拟合效果和预测效果, 并对其稳定性进行检验, 发现 EAPC 模型更适合于拟合和预测中国的人口死亡率。肖鸿民等(2020) [14]分别运用 ARIMA 方法和双随机过程对 LC 模型的时间项进行预测, 通过比较它们的预测结果表明, 经典 ARIMA 方法更适合目前中国的人口死亡率预测。近些年来, 随着机器学习的不断深入, 许多学者开始将机器学习方法应用到死亡率的预测研究中, 而且发现其明显地提高了死亡率的预测效果。Deprez 等(2017) [15]发现了机器学习算法可以有效评估标准的动态死亡率预测模型(LC 模型和 RH 模型)的拟合优度。Hainaut (2018) [16]使用自编码器将死亡率中蕴含的信息进行编码和解码, 并通过带漂移项的随机游走模型进行预测, 其死亡率预测性能要优于 LC 模型和 RH 模型。Levan-tesì 等(2019) [17]基于决策树、随机森林、梯度提升树等树方法对死亡率数据进行拟合和预测, 其效果显著优于 LC 模型、RH 模型和 Plat 模型。陈宁(2019) [18]在 Hainaut (2018)提出的神经网络死亡率模型的基础上, 根据我国人口死亡率数据的特征, 对模型的估计方法和预测方法进行调整, 构建了适合我国人口的神经网络死亡率模型。在模型的估计方法上, 文章采用降噪自编码神经网络来提取潜在时间因子, 在模型的预测方法上, 应用 ARIMA 模型对潜在时间因子进行外推。通过将模型的预测结果与 LC 模型的预测结果进行对比, 对模型的预测效果进行分析。结果表明, 文章建立的神经网络死亡率模型的预测结果整体上优于 LC 模型的预测结果, 该模型可以较好地预测我国人口死亡率。张连增等(2020) [19]应用稀疏自编码器, 建立死亡率的神经网络模型, 使模型通过训练能够自行学习到死亡率的潜在特征, 并将模型的拟合结果和预测结果与传统 LC 模型进行比较分析。结果显示, 自编码器对死亡率的拟合效果与模型相近, 而预测效果显著优于 LC 模型, 说明自编码器显著提高了对死亡率的预测性能。

目前为止, 大多死亡率的预测部分都是基于传统的 ARIMA 方法进行的, Wei Hong 等(2020) [20]提出 LC-ANN 模型, 即将 ANN 用于死亡率模型 LC 的预测过程中, 分别对马来西亚男性、女性死亡率进行预测, 结果表明 LC-ANN 模型有着良好的预测能力。本文以我国人口死亡率数据为例, 以经典死亡率模型(LC 模型)为出发点, 对模型进行参数估计, 考虑其时间项受非线性因素的影响, 尝试将 LSTM 网络融进基于循环神经网络的死亡率的预测分析过程, 得到死亡率预测值, 与传统的 ARIMA 预测方法进行比较。

2. 随机死亡率模型

2.1. 模型简介

LC 模型假设对数死亡率与时间呈现关联性。以 $m_{x,t}$ 表示 x 岁的人在 t 年内的中心死亡率, 在传统的 LC 模型中, 中心死亡率 $m_{x,t}$ 的对数与年龄、时间的关系可用下式表示:

$$\ln m_{x,t} = \alpha_x + \beta_x k_t + \epsilon_{x,t} \quad (1)$$

其中 $m_{x,t}$ 表示 t 年龄为 x 岁群体的中心死亡率; $\alpha_x (x=0,1,2,\dots)$ 为依赖年龄因子的参数, 表示的是 x 岁人口对数死亡率的平均数; $\beta_x (x=0,1,2,\dots)$ 表示年龄因子对时间的敏感度; k_t 反映了死亡率随时间 t 的变化程度, 通常称为死亡指数; $\epsilon_{x,t}$ 为随机误差项。

LC 模型对死亡率的拟合和预测分为两步, 第一步基于死亡率数据, 对参数 α_x 、 β_x 和 k_t 运用¹极大似然估计方法进行估计。第二步对时间因子 k_t 建立时间序列模型, 并预测未来的时间因子, 将预测值代入 LC 模型中, 并根据 α_x 和 β_x 的估计值, 得到预测的死亡率数值。

¹极大似然估计法的一个优点是不需要对 k_t 的估计值进行修正, 因为该方法直接假设死亡数服从泊松分布, 而 SVD (Singular Value Decomposition)法求出的 k_t 的估计值需要进一步修正。

2.2. 参数估计

对上述随机死亡率模型进行参数估计时, 都假定死亡人数 $D_{x,t}$, 服从参数为 $\lambda_{x,t}$ 的泊松分布, $\lambda_{x,t} = m_{x,t} E_{x,t}$, 即

$$D_{x,t} \sim \text{Poisson}(m_{x,t} E_{x,t}) \quad (2)$$

其中, $m_{x,t} = \exp(\alpha_x + \beta_x k_t)$, $D_{x,t}$ 表示年龄为 x 、时间为 t 的死亡人数, $E_{x,t}$ 表示年龄为 x 、时间为 t 的暴露人数。

根据以上假设, 对于任一年龄 x 、时刻 t , 似然函数(不考虑下标)和对数似然函数可以表示为

$$L(d; \lambda) = \frac{\lambda^d}{d!} e^{-\lambda} \quad (3)$$

$$l = \ln L(d; \lambda) = d \ln(\lambda) - \lambda - \ln(d!) \quad (4)$$

若假设每一年龄的死亡相对独立, 则总体对数似然函数表示为:

$$\begin{aligned} l(\alpha, \beta, k) &= \sum_{x,t} [d_{x,t} \ln(\lambda_{x,t}) - \lambda_{x,t} - \ln(d!)] \\ &= \sum_{x,t} [d_{x,t} (\alpha_x + \beta_x k_t) - E_{x,t} \exp(\alpha_x + \beta_x k_t)] + c \end{aligned} \quad (5)$$

对于一个给定的模型, 用 θ 来代表待估计的参数向量, 运用牛顿迭代法得到各参数的估计值, 牛顿迭代公式如下:

$$\hat{\theta}^{(n+1)} = \hat{\theta}^{(n)} - \frac{\partial L^{(n)} / \partial \theta}{\partial^2 L^{(n)} / \partial \theta^2} \quad (6)$$

其中, $\hat{\theta}^{(n)}$ 是第 n 次迭代参数, $L^{(n)} = L[\hat{\theta}^{(n)}]$ 。

因此 LC 模型参数迭代具体内容为:

$$\hat{\alpha}_x^{(n+1)} = \hat{\alpha}_x^{(n)} - \frac{\sum_t D_{x,t} - \hat{D}_{x,t}^{(n)}}{-\sum_t \hat{D}_{x,t}^{(n)}}, \hat{\beta}_x^{(n+1)} = \hat{\beta}_x^{(n)}, \hat{k}_t^{(n+1)} = \hat{k}_t^{(n)} \quad (7)$$

$$\hat{k}_t^{(n+2)} = \hat{k}_t^{(n+1)} - \frac{\sum_x (D_{x,t} - \hat{D}_{x,t}^{(n+1)}) \hat{\beta}_x^{(n+1)}}{-\sum_x \hat{D}_{x,t}^{(n+1)} (\hat{\beta}_x^{(n+1)})^2}, \hat{\alpha}_x^{(n+2)} = \hat{\alpha}_x^{(n+1)}, \hat{\beta}_x^{(n+2)} = \hat{\beta}_x^{(n+1)} \quad (8)$$

$$\hat{\beta}_x^{(n+3)} = \hat{\beta}_x^{(n+2)} - \frac{\sum_t (D_{x,t} - \hat{D}_{x,t}^{(n+2)}) \hat{k}_t^{(n+2)}}{-\sum_t \hat{D}_{x,t}^{(n+2)} (\hat{k}_t^{(n+2)})^2}, \hat{\alpha}_x^{(n+3)} = \hat{\alpha}_x^{(n+2)}, \hat{k}_t^{(n+3)} = \hat{k}_t^{(n+2)} \quad (9)$$

其中 $\hat{D}_{x,t}^{(n)} = E_{x,t} \exp(\hat{\alpha}_x^{(n)} + \hat{\beta}_x^{(n)} \hat{k}_t^{(n)})$ 或 $\hat{D}_{x,t}^{(n)}$ 为第 n 步迭代得到的死亡数估计值。利用更新的参数重新计算对数似然函数, 当参数更新使对数似然函数收敛(即增加量很小, 通常小于 10^{-10})时停止迭代, 得到参数估计值。

3. LSTM 神经网络

3.1. 理论准备

LSTM (Long Short Term Memory, LSTM)网络[21] [22], 是一种 RNN (Recurrent Neural Network, 循

神经网络)特殊的类型, 可以学习长期依赖序列的信息, 通过训练时在“时间上的反向传播”来克服梯度消失的问题。LSTM 网络可以用来构建大规模的递归神经网络来处理机器学习中复杂的序列问题, 与此同时, 得到不错的效果。

与 RNN [23]不同的是, 他在算法中加入了一个判断信息有用与否的“处理器”, 这个处理器的作用结构被称作是激活单元(cell)。一个 cell 中被放置了三扇门, 分别叫做输入门(Input Gate)、遗忘门(Forget Gate)、输出门(Output Gate)。一个信息进入 LSTM 网络中, 只有符合信息的才会留下, 不符合的信息则通过遗忘门被遗忘。有了这种机制的存在, 理论上减少了梯度爆炸(消失)的风险。

3.2. 模型构建

时间序列预测相对是较困难的一类问题。最简单明了的思路就是寻找当前和过去数据 (X_t, X_{t-1}, \dots) 与未来数据 (X_{t+1}) 之间的关系, 一般情况下将这种关系表示成为一个回归问题。本文应用 R 软件中的 keras 包, 以 LSTM 网络回归为出发点, 通过 keras 包中的 keras_model_sequential, layer_lstm, layer_dense 函数建立神经网络, compile 函数编译模型的优化目标, 并用 fit 函数进行拟合。

4. 预测结果分析

4.1. 数据来源

本文采用的数据来源于 1995~2017 年的《中国人口统计年鉴》, 选取连续 23 年的中国人口 0~90 岁死亡率数据。大部分日历年死亡率数据的末组都为 90+ (90 岁及以上用 90+表示), 因此本文的年龄上限设为 90。

4.2. 参数估计

选取上述数据 1995~2013 年的死亡率数据, 结合(7)、(8)、(9)式对 LC 模型进行参数估计, 结果如下表, 见表 1。从图 1 中可以看出, 参数 α_x 表示 x 岁人口对数死亡率的平均水平, 随着年龄的增长, 其值呈现了先下降后上升的趋势; β_x 作为年龄因子对时间的敏感度, 其低年龄的值较高, 高年龄的值逐渐趋向 0, 表明低年龄人群对死亡率时间因子的变化更加敏感, 高年龄组人口死亡率的时间敏感度较小; k_t 的估计值随着时间的推移而逐渐降低。

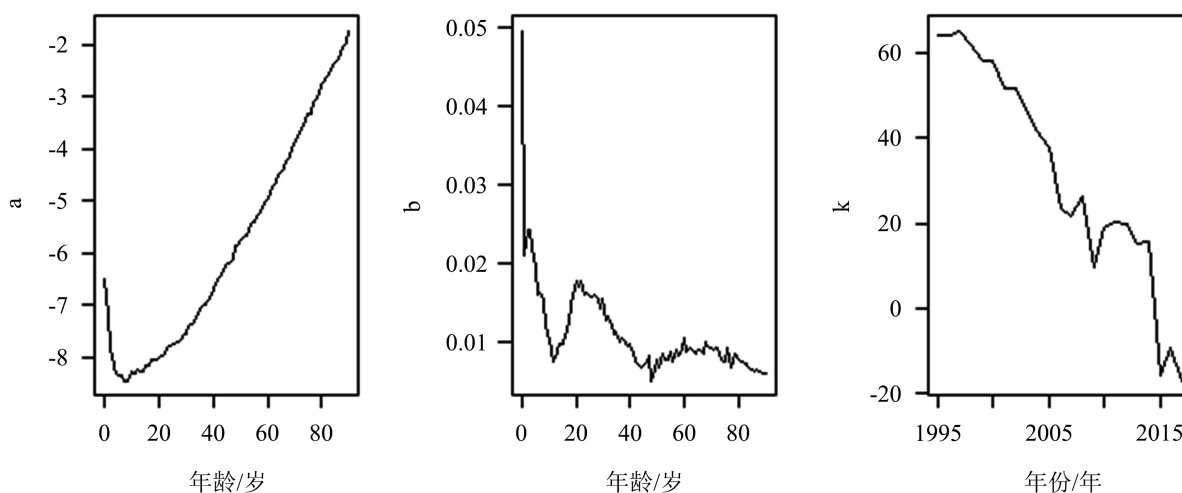


Figure 1. Parameters estimation
图 1. 参数估计

Table 1. Estimates of $\hat{\alpha}_x$ and $\hat{\beta}_x$ **表 1.** $\hat{\alpha}_x$ 和 $\hat{\beta}_x$ 的估计值

age	$\hat{\alpha}_x$	$\hat{\beta}_x$	age	$\hat{\alpha}_x$	$\hat{\beta}_x$	age	$\hat{\alpha}_x$	$\hat{\beta}_x$
0	-6.485755	0.049524	31	-7.409944	0.012716	62	-4.738761	0.009012
1	-7.198779	0.020768	32	-7.361465	0.013027	62	-4.667308	0.008762
2	-7.830775	0.024110	33	-7.323007	0.012140	63	-4.425088	0.008303
3	-8.154655	0.024250	34	-7.171831	0.010912	64	-4.387482	0.008935
4	-8.315216	0.021919	35	-7.093256	0.011141	65	-4.221649	0.008949
5	-8.365536	0.019267	36	-7.032488	0.009878	66	-4.169830	0.008340
6	-8.358963	0.015974	37	-6.979232	0.010494	67	-4.000924	0.009830
7	-8.461958	0.016142	38	-6.908745	0.010090	68	-3.845957	0.009423
8	-8.461274	0.015274	39	-6.790529	0.009456	69	-3.793029	0.009199
9	-8.397840	0.011985	40	-6.675657	0.009538	70	-3.648517	0.008879
10	-8.286131	0.010071	41	-6.629195	0.008605	71	-3.555675	0.009007
11	-8.307184	0.008829	42	-6.449000	0.007348	72	-3.436063	0.008381
12	-8.253405	0.007470	43	-6.412013	0.007117	73	-3.328049	0.008381
13	-8.299446	0.008700	44	-6.312303	0.006716	74	-3.337263	0.007712
14	-8.284689	0.009693	45	-6.201040	0.006787	75	-3.105504	0.007461
15	-8.153931	0.009627	46	-6.188423	0.007416	76	-3.022168	0.009118
16	-8.150109	0.010607	47	-6.134400	0.008166	77	-3.105504	0.006511
17	-8.058737	0.011165	48	-5.872851	0.004936	78	-2.748997	0.007347
18	-8.062973	0.014415	49	-5.847377	0.006449	79	-2.693988	0.008332
19	-8.051155	0.015831	50	-5.765416	0.007645	80	-2.600064	0.007584
20	-8.007657	0.017679	51	-5.713672	0.006740	81	-2.510195	0.007328
21	-7.973424	0.016873	52	-5.688740	0.008439	82	-2.395216	0.007403
22	-7.936838	0.017557	53	-5.576178	0.007652	83	-2.327588	0.007149
23	-7.831353	0.015875	54	-5.444280	0.007531	84	-2.249628	0.006536
24	-7.786893	0.016131	55	-5.424943	0.008543	85	-2.162488	0.006468
25	-7.746648	0.015897	56	-5.317418	0.007452	86	-2.057567	0.006176
26	-7.759412	0.015594	57	-5.267246	0.008868	87	-1.964543	0.006264
27	-7.726829	0.015981	58	-5.147719	0.008245	88	-2.162488	0.006029
28	-7.689415	0.015375	59	-5.039057	0.008870	89	-1.964543	0.005800
29	-7.557582	0.014048	60	-4.958702	0.010389	90	-1.720298	0.005833
30	-7.556472	0.015389	61	-4.834724	0.008555			

4.3. 死亡率预测

对得到的时间项估计值进行最大-最小特征缩放, 见表 2。划分数据集, 取 1995~2011 年数据作为训练集, 2012~2017 年数据集为测试集, 带入前文搭建好的神经网络模型进行训练、拟合和预测, 预测结束后将预测数据进行反归一化, 得到预测值, 见表 3。

4.4. 比较分析

见图 2 所示, 通过与实际死亡率数据以及传统的 ARIMA 方法预测值比较, 基于循环神经网络的 LSTM 学习机在预测中国人口死亡率的应用中有良好的效果。下面我们尝试适当地改变 LSTM 学习机的网络结构, 寻求更适合中国人口死亡率预测的神经网络结构, 同样得到新的死亡率预测值, 与前面的两次结果作比较。

见图 3 所示, 在适当的改变 LSTM 网络层单元数后, 预测值(RNN-LSTM2)明显优于前文所提到的 ARIMA 方法得到的预测值与起初的神经网络预测(RNN-LSTM1), 故充分展示了神经网络在死亡率预测中的优越性。

Table 2. Time-term estimated value after feature scaling

表 2. 经过特征缩放后的时间项估计值

year	k_t	year	k_t	year	k_t
1995	0.98479325	2003	0.77086595	2011	0.45249095
1996	0.98469093	2004	0.70996925	2012	0.44552309
1997	1.00000000	2005	0.66548162	2013	0.39046947
1998	0.96067694	2006	0.49316280	2014	0.39483149
1999	0.91298239	2007	0.47151582	2015	0.01595222
2000	0.91137558	2008	0.52402043	2016	0.09414175
2001	0.83687991	2009	0.32367418	2017	0.00000000
2002	0.83754272	2010	0.43880770		

Table 3. Mortality estimated value (Unit: %)

表 3. 死亡率估计值(单位: %)

Age	year			Age	year		
	2015	2016	2017		2015	2016	2017
0	0.610	0.315	0.106	47	0.272	0.244	0.249
1	0.133	0.101	0.060	48	0.323	0.302	0.306
2	0.078	0.056	0.043	49	0.345	0.317	0.322
3	0.057	0.041	0.035	50	0.388	0.350	0.357
4	0.045	0.033	0.032	51	0.398	0.364	0.370
5	0.040	0.030	0.030	52	0.428	0.383	0.391
6	0.037	0.029	0.027	53	0.469	0.423	0.432
7	0.033	0.026	0.027	54	0.533	0.482	0.491
8	0.032	0.026	0.027	55	0.559	0.499	0.510

Continued

9	0.031	0.026	0.029	56	0.604	0.547	0.557
10	0.033	0.029	0.028	57	0.661	0.587	0.600
11	0.031	0.028	0.029	58	0.732	0.656	0.669
12	0.032	0.029	0.028	59	0.830	0.738	0.754
13	0.031	0.028	0.029	60	0.939	0.818	0.839
14	0.033	0.029	0.033	61	1.010	0.901	0.921
15	0.037	0.033	0.034	62	1.126	0.998	1.021
16	0.038	0.033	0.038	63	1.201	1.068	1.092
17	0.043	0.037	0.040	64	1.396	1.250	1.276
18	0.047	0.038	0.041	65	1.537	1.365	1.396
19	0.049	0.040	0.045	66	1.597	1.417	1.450
20	0.054	0.043	0.046	67	1.853	1.658	1.694
21	0.055	0.044	0.048	68	2.035	1.785	1.830
22	0.058	0.046	0.052	69	2.382	2.101	2.152
23	0.061	0.050	0.054	70	2.764	2.445	2.503
24	0.065	0.052	0.056	71	2.888	2.566	2.624
25	0.067	0.054	0.055	72	3.349	2.971	3.039
26	0.066	0.053	0.058	73	3.611	3.230	3.299
27	0.068	0.055	0.059	74	3.995	3.605	3.676
28	0.070	0.057	0.066	75	4.419	4.001	4.077
29	0.077	0.064	0.068	76	4.587	4.063	4.157
30	0.080	0.065	0.075	77	5.376	4.930	5.011
31	0.086	0.072	0.079	78	5.981	5.424	5.526
32	0.091	0.076	0.081	79	6.688	5.986	6.113
33	0.092	0.078	0.092	80	7.913	7.153	7.291
34	0.104	0.090	0.100	81	8.301	7.529	7.670
35	0.113	0.097	0.104	82	9.137	8.280	8.436
36	0.116	0.102	0.111	83	9.926	9.025	9.189
37	0.124	0.108	0.118	84	10.946	10.034	10.201
38	0.132	0.115	0.132	85	11.689	10.725	10.902
39	0.146	0.129	0.148	86	12534	11.545	11.726
40	0.164	0.145	0.153	87	13709	12.613	12.813
41	0.168	0.149	0.179	88	15.126	13.960	14.173
42	0.194	0.176	0.185	89	16.494	15.269	15.494
43	0.200	0.182	0.203	90	21.077	19.503	19.792
44	0.218	0.200	0.227				
45	0.245	0.224	0.233				
46	0.252	0.228	0.233				

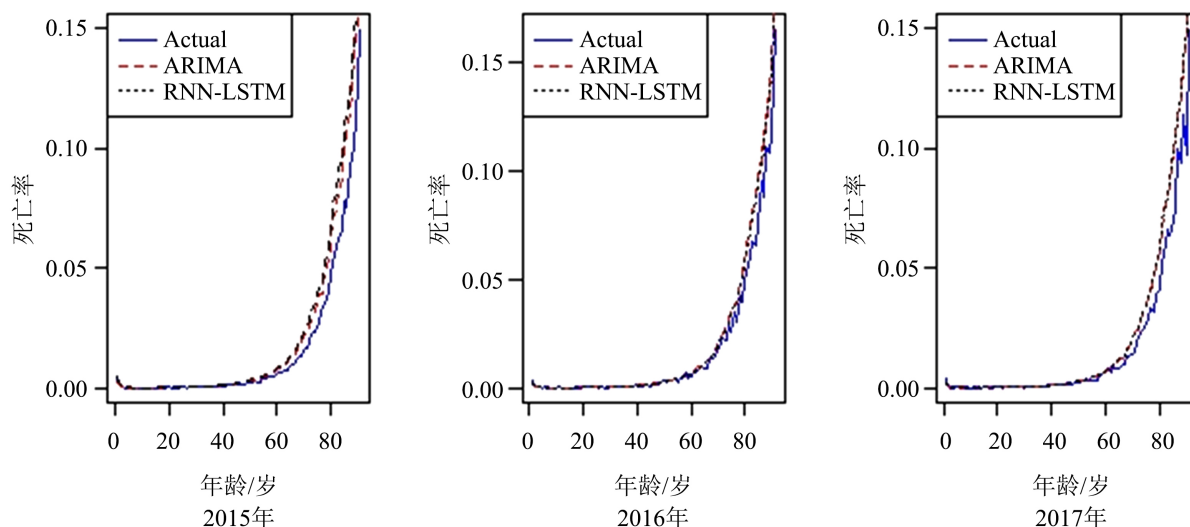


Figure 2. Comparison of actual mortality with ARIMA and RNN-LSTM estimated value

图 2. 实际死亡率与 ARIMA 及 RNN-LSTM 估计值的比较

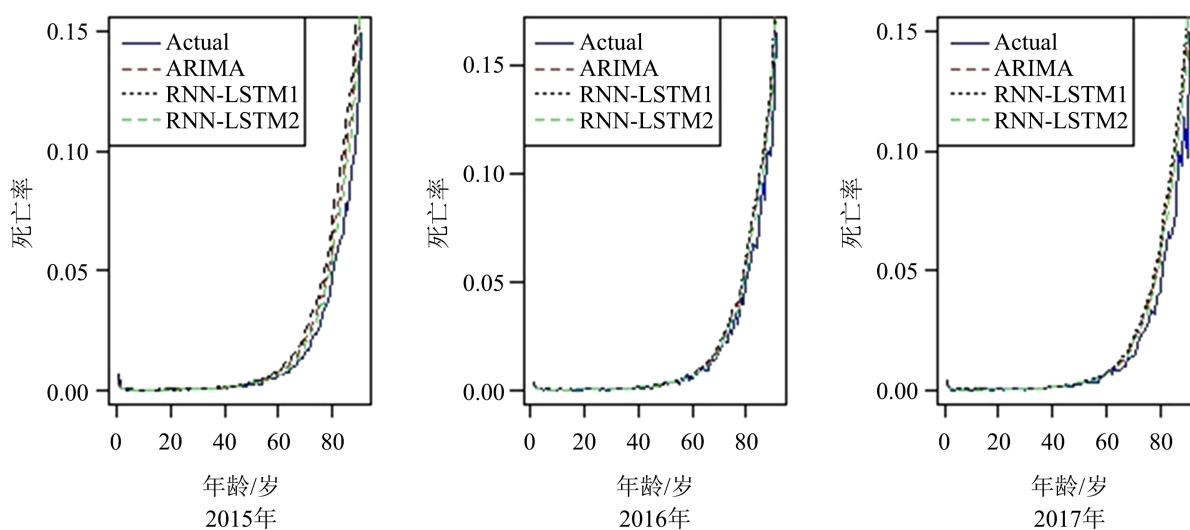


Figure 3. Comparison between the predicted value of changing the number of cells in the LSTM network layer and the predicted value of the ARIMA method

图 3. 改变 LSTM 网络层单元数预测值与真实值及 ARIMA 方法预测值的比较

5. 结论

死亡率的下降和预期寿命的延长显然成为我国人口发展的主要趋势, 由此带来的长寿风险不容忽视, 特别是在养老金体系中, 基本养老保险、企业补充养老保险以及商业养老保险都需重视未来死亡率发展趋势, 并采取相应的措施来应对长寿风险。以往的死亡率时间项大多都是运用传统的 ARIMA 方法进行线性外推, 考虑到非线性因素的影响, 本文应用基于 Keras 的 LSTM 循环神经网络预测模型, 不仅解决了 RNN 神经网络中存在的梯度消失和梯度爆炸问题, 同时也解决了很多古典线性方法难以适应多个变量或多输入的时间序列预测问题, 因此该模型能够应用于死亡率的预测。本文通过对我国数据的拟合预测, 可以核实基于循环神经网络的 LSTM 学习机在预测中国人口死亡率的应用中有良好的效果, 并且在适当的改变网络结构后, 发现得到的死亡率预测值明显优于传统的 ARIMA 方法得到的预测值。

基金项目

国家自然科学基金资助项目(12061066); 甘肃省自然科学基金(20JR5RA528)。

参考文献

- [1] Lee, R.D. and Carter, L.R. (1992) Modeling and Forecasting US Mortality. *Journal of American Statistical Association*, **87**, 659-675. <https://doi.org/10.2307/2290201>
- [2] Renshaw, A.E. and Haberman, S. (2006) A Cohort-Based Extension to the Lee-Carter Model for Mortality Reduction Factors. *Insurance: Mathematics and Economics*, **38**, 556-570. <https://doi.org/10.1016/j.insmatheco.2005.12.001>
- [3] Cairns, A.J.G., Blake, D., Dowd, K., et al. (2011) Mortality Density Forecasting: An Analysis of Six Stochastic Mortality Models. *Insurance: Mathematics and Economics*, **48**, 355-367. <https://doi.org/10.1016/j.insmatheco.2010.12.005>
- [4] Li, N., Lee, R. and Gerland, P. (2013) Extending the Lee-Carter Method to Model the Rotation of Age Patterns of Mortality Decline for Long-Term Projections. *Demography*, **50**, 2037-2051. <https://doi.org/10.1007/s13524-013-0232-2>
- [5] Enchev, V., Kleinow, T. and Cairns, A.J.G. (2017) Multi-Population Mortality Models: Fitting, Forecasting and Comparisons. *Scandinavian Actuarial Journal*, **2017**, 319-342. <https://doi.org/10.1080/03461238.2015.1133450>
- [6] Li, J. and Liu, J. (2019) A Logistic Two-Population Mortality Projection Model for Modelling Mortality at Advanced Ages for Both Sexes. *Scandinavian Actuarial Journal*, **2019**, 97-112. <https://doi.org/10.1080/03461238.2018.1511464>
- [7] 李志生, 刘恒甲. Lee-Carter 的估计和应用——基于中国数据分析[J]. 中国人口科学, 2010(3): 47-56.
- [8] 黄顺林, 王晓军. 加入出生年效应的死亡率预测及其在年金系数估计中的应用[J]. 统计与信息论, 2010(5): 81-86.
- [9] 韩猛, 王晓军. Lee-Carter 模型在中国城市人口死亡率预测中的应用与改进[J]. 保险研究, 2010(10): 3-9.
- [10] 曾燕, 陈曦, 邓颖璐. 创新的动态人口死亡率预测及其应用[J]. 系统工程理论与实践, 2016, 36(7): 1710-1718.
- [11] 曹园. 基于 Lee-Carter 模型的我国死亡率预测[J]. 统计与决策, 2018(9): 32-36.
- [12] 樊毅, 张宁. 基于全人口死亡率数据的随机死亡率模型拟合效果比较[J]. 统计与决策, 2018(23): 33-37.
- [13] 王晓军, 陆倩. 高龄人口死亡率预测模型的选择与比较[J]. 保险研究, 2019(3): 82-102.
- [14] 马海飞, 肖鸿民, 赵弘宇. 一种新的死亡率模型及基于中国人口数据的比较分析[J]. 经济数学, 2020, 37(3): 99-106.
- [15] 肖鸿民, 马海飞, 康彦玲. 两种死亡率预测方法的比较[J]. 统计与决策, 2020, 36(23): 5-8.
- [16] Deprez, P., Shevchenko, P.V. and Wüthrich, M.V. (2017) Machine Learning Techniques for Mortality Modeling. *European Actuarial Journal*, **7**, 337-352. <https://doi.org/10.1007/s13385-017-0152-4>
- [17] Hainaut, D. (2018) A Neural-Network Analyzer for Mortality Forecasting. *ASTIN Bulletin: The Journal of the IAA*, **48**, 481-508. <https://doi.org/10.1017/asb.2017.45>
- [18] 陈宁. 基于神经网络死亡率模型的我国死亡率预测及应用[D]: [硕士学位论文]. 天津: 天津财经大学, 2019.
- [19] 张连增, 申晴, 丁宁. 自编码器在死亡率预测中的应用研究[J]. 保险研究, 2020(7): 83-93.
- [20] Hong, W.H., et al. (2020) Forecasting Mortality Rates Using Hybrid Lee-Carter Model, Artificial Neural Network and Random Forest. *Complex & Intelligent Systems*, **7**, 163-189. <https://doi.org/10.1007/s40747-020-00185-w>
- [21] Giuseppe Ciaburro, Balaji Venkateswaran. R 语言实现[M]. 李洪成, 译. 北京: 机械工业出版社, 2018.
- [22] 贾明珠. 基于机器学习的时间序列分析方法研究与应用[D]: [硕士学位论文]. 西安: 西安科技大学, 2020.
- [23] 杨丽, 吴雨茜, 王俊丽, 刘义理. 循环神经网络研究综述[J]. 计算机应用, 2018, 38(S2): 1-6+26.