

Prediction of Population Aging in Chengdu: A Study Based on Optimized Grey GM(1,1) Model

Miao Yang

L.A.S Department of CDUESTC, Chengdu College of University of Electronic Science and Technology of China,
Chengdu Sichuan

Email: yangmiao603@163.com

Received: Oct. 2nd, 2017; accepted: Oct. 16th, 2017; published: Oct. 23rd, 2017

Abstract

On the basis of data of 2009~2016 population aging, the paper uses Optimized Grey GM(1,1) model and metabolic GM(1,1) model, with testing and programming calculation of MATLAB software, to predict population aged 60 and older between 2017~2022 of Chengdu City. Model accuracy test shows that the model accuracy is good. The model prediction results are high reliability and believability. The model is suitable for the long-term prediction of population.

Keywords

Optimized Grey GM(1,1) Model, Population Aging, Metabolic GM(1,1) Model

基于优化GM(1,1)灰色模型预测成都市 老年人口

杨 淼

电子科技大学成都学院文理系, 四川 成都

Email: yangmiao603@163.com

收稿日期: 2017年10月2日; 录用日期: 2017年10月16日; 发布日期: 2017年10月23日

摘 要

以2009~2016年成都市老年人口数据资料为依据, 通过优化的GM(1,1)模型和新陈代谢GM(1,1)模型,

文章引用: 杨淼. 基于优化 GM(1,1)灰色模型预测成都市老年人口[J]. 统计学与应用, 2017, 6(4): 396-401.

DOI: 10.12677/sa.2017.64045

借助Matlab软件编程计算和检验,用该模型预测了成都市2017~2022年老龄人口规模,并对预测结果进行了分析。模型精度检验表明模型精度为较好,模型预测结果可靠性高、可信度强,模型适合人口中长期预测。

关键词

优化GM(1,1)模型, 老年人口, 新陈代谢模型

Copyright © 2017 by author and Hans Publishers Inc.

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

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



Open Access

1. 引言

人口老龄化是指总人口中因年轻人口数量减少、年老人口数量增加而导致的老年人口比例相应增长的动态。国际上通常看法是,当一个国家或地区 60 岁以上老年人口占人口总数的 10%, 或 65 岁以上老年人口占人口总数的 7%, 即意味着这个国家或地区的人口处于老龄化社会。通过对影响老龄人口规模变化因素的历史资料进行统计分析,探讨其在时间上的变化规律,可以得出老龄人口规模变化的长期趋势,从而对它的未来变化进行预测。灰色预测法是一种对既含有已知信息又含有不确定因素的系统进行预测的方法,它的特点是所需信息量少,不仅能够将无序离散的原始序列转化为有序序列,而且预测精度高,能够保持原系统的特征,较好地反映系统地实际情况。本文运用优化的 GM(1,1)模型[1] [2] [3]和改进的新陈代谢 GM(1,1)模型对成都市老龄人口规模进行预测,取得了良好地效果。

2. 数学模型简介

2.1. GM(1,1)模型

GM(1,1)模型是基于累加生成地数列预测模型,建立地步骤为:

1) $x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)$ 是所要预测地某项指标地原始数据。对原始数据作一次累加生成处理,即 $x^{(1)}(n) = \sum_{i=1}^n x^{(0)}(i)$, 得到一个新地数列。这个新的数列与原始数列相比,其随机性程度大大弱化,平稳性大大增加。

2) 将新数列的变化趋势近似地用微分方程描述, $\frac{dx^{(1)}}{dt} + \alpha x^{(1)} = \mu$, 其中, α, μ 为辨识参数。辨识参数通过最小二乘法拟合得到, $\begin{bmatrix} \alpha \\ \mu \end{bmatrix} = (B^T B)^{-1} B^T Y_n$ 。

3) 构造矩阵。 $Y_n = [x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)]^T$, B 为构造数据矩阵

$$B = \begin{bmatrix} -\frac{1}{2}[x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2) + x^{(1)}(3)] & 1 \\ \vdots & \vdots \\ -\frac{1}{2}[x^{(1)}(n-1) + x^{(1)}(n)] & 1 \end{bmatrix} \quad (1)$$

4) 求出预测模型

$$\hat{x}^{(1)}(t+1) = \left[x^{(0)}(1) - \frac{\mu}{\alpha} \right] e^{-\alpha t} + \frac{\mu}{\alpha}. \quad (2)$$

5) 进行一次数据累减, 得到预测方程:

$$\hat{x}^{(0)}(t+1) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t). \quad (3)$$

2.2. 优化 GM(1,1)模型[4]

由上述的 GM(1,1)模型的模拟, 可知预测精度取决于常数 α, μ , 而 α, μ 的值依赖于原始序列和背景值 $z^{(1)}(t) = \frac{1}{2}(x^{(1)}(t) + x^{(1)}(t-1))$ 的构造形式。即背景值的构造公式是导致模拟误差及 GM(1,1)模型的适应性的关键因素之一。

故我们为了尽量减小这种误差, 设

$$z^{(1)}(t) = \int_{t-1}^t x^{(1)}(t) dt. \quad (4)$$

设 $x^{(1)}(t) = Be^{At}$, 其中 A, B 为待定常数, 且满足 $x^{(1)}(t) = Be^{At}, (t=1, 2, \dots, n)$, 将 $x^{(1)}(t) = Be^{At}$ 代入(4)得

$$z^{(1)}(t) = \int_{t-1}^t Be^{At} dt = \frac{1}{A} [Be^{At} - Be^{A(t-1)}] = \frac{1}{A} [x^{(1)}(t) - x^{(1)}(t-1)] \quad (5)$$

又由 $\frac{x^{(1)}(t)}{x^{(1)}(t-1)} = \frac{Be^{At}}{Be^{A(t-1)}} = e^A$ 得

$$A = \ln x^{(1)}(t) - \ln x^{(1)}(t-1). \quad (6)$$

将(6)代入(5)得

$$z^{(1)}(t) = \frac{x^{(1)}(t) - x^{(1)}(t-1)}{\ln x^{(1)}(t) - \ln x^{(1)}(t-1)}, k = 2, 3, \dots, n \quad (7)$$

用(7)替换原来的背景值公式, 由此得到优化的 GM(1,1)模型。

2.3. 新陈代谢 GM(1,1)模型[5]

在普通 GM(1,1)模型的基础上, 我们建立一种新模型叫做新陈代谢 GM(1,1)模型, 其方法是把普通 GM(1,1)模型预测的最新数据 $\hat{x}^{(0)}(t+1)$ 加入到原始数据序列 $x^{(0)}$ 中, 然后去掉最老的数据 $x^{(0)}(1)$, 为了使数据序列的维度保持不变, 接着用更新的数据序列 $x^{(0)} = \{x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(t+1)\}$ 再次进行普通 GM(1,1)模型的建立与检验, 预测出新数据 $x^{(0)}(t+2)$, 然后再将 $x^{(0)}(t+2)$ 加入到 $x^{(0)} = \{x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(t+1)\}$, 再去掉 $x^{(0)}(2)$, 形成新数列, 重复普通 GM(1,1)模型的预测与检验。往复循环, 数据更迭, 达到需要的预测目标为止, 这种模型称为新陈代谢 GM(1,1)模型。

新陈代谢 GM(1,1)模型的原始数据序列的数据不应太少, 若原始数据序列的数据太少, 则有可能得到关联度比较低, 模型不能通过检验, 就不能作为预测方法来预测, 特别是预测中长期的数据; 当然, 原始数据的数据也不是越多越好, 在实际的数据预测时, 可以利用普通 GM(1,1)模型, 选取不同数据个数的原始数据序列逐个进行检验, 找到检验精度最高的那组原始数据序列作为原始数据序列, 由此确定 GM(1,1)模型的维度, 进行新陈代谢 GM(1,1)模型的模拟。

3. 实例分析

3.1. 预测模型的建立

笔者收集了 2011~2016 年成都市老年人口的统计资料(见表 1), 对 2017~2022 年成都市老年人口做出预测。

由表 1 可得成都市老年人口的原始时间序列为 $x^{(0)}(t)=[197\ 203\ 214\ 223\ 236\ 249\ 260\ 300]$, 由一次累加数据序列 $x^{(1)}(t)=[197\ 400\ 614\ 837\ 1073\ 1322\ 1582\ 1882]$, 由(1)式可得

$$B^T = \begin{bmatrix} -256.6 & -499.4 & -719.8 & -950.1 & -1193.2 & -1448.1 & -1727.7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$Y_n = [400, 614, 837, 1073, 1322, 1582, 1882]$$

由此可得 $\begin{bmatrix} \alpha \\ \mu \end{bmatrix} = \begin{bmatrix} -0.0616 \\ 180.6150 \end{bmatrix}$, 代入(2), (3)可得预测模型

$$\begin{cases} \hat{x}^{(1)}(t+1) = 3129.1e^{0.0616t} - 2932.1 \\ \hat{x}^0(t+1) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t) \end{cases}$$

3.2. 模型检验

为了最大限度的减小误差取得更好的预测效果, 建立数据处理模型后需通过相关项的精度检验, 否则不能进行预测, 检验方法包括以下三种:

1) 残差检验。

逐点检验模型预测还原值与实际值的残差[6] [7]。绝对残差序列: $\Delta^{(0)}(k) = |x^{(0)}(k) - \hat{x}^{(0)}(k)|$, 平均相对残差: $\bar{\Phi} = \frac{1}{k} \sum_{i=1}^k \varphi_i$, 其中 $\varphi_i = \frac{\Delta^{(0)}(k)}{x^{(0)}(k)}$, $(i=1, 2, 3, \dots, n)$ 。

在实际检验中, 我们给定一个 α 值, 当 $\bar{\Phi} < \alpha$ 且 $\varphi_i < \alpha$ 成立时, 称模型为残差合格模型。此篇文章中我们定义 α 的值如下: 当 α 取 0.02 时, 模型精度等级为优, 当 α 取 0.05 时, 模型精度为合格, 当 α 取 0.1 时, 模型精度为勉强合格。

2) 关联度检验。

检验模型序列曲线与原始序列曲线的相似程度[6] [7]。

令关联度 $R = \frac{l}{n} \sum_{i=1}^n \xi_i(t)$, 令关联系数 $\xi_i = \frac{\Delta_{\min} + k\Delta_{\max}}{\Delta_i + k\Delta_{\max}}$, 其中 $\Delta_{\min} = \min \Delta_i$, $\Delta_{\max} = \max \Delta_i$, $\Delta_i = [x_i^{(1)}(t) - x_i^{(0)}(t)]$, 通过关联度检验的标准就是关联系数大于 0.6, 当关联度检验通过时, 我们给出的预测模型便是满意。

3) 后验差检验。检验残差分布的统计特性[6] [7]。

由原始数据序列和绝对误差序列计算得原始数据序列和绝对误差序列得标准差分别为:

Table 1. The number of population aging in Chengdu (unit: ten thousand people)

表 1. 成都市老年人口数量(单位: 万人)

年份	2009	2010	2011	2012	2013	2014	2015	2016
人口	197	203	214	223	236	249	260	300

$$S_1 = \sqrt{\frac{\sum (x^{(0)}(t) - \hat{x}^{(0)}(t))^2}{n-1}}$$

$$S_2 = \sqrt{\frac{\sum (\Delta^{(0)}(t) - \hat{\Delta}^{(0)}(t))^2}{n-1}}$$

由此计算得方差比 $C = \frac{S_2}{S_1}$ ，小误差概率 $P = \left\{ \left| \Delta^0(t) - \bar{\Delta}^0 \right| < 0.6745S_1 \right\}$ ，一般情况下，模型预测精度检验等级标准见表 2，模型精度等级为较好时，则 $p > 0.95$ ， $C < 0.35$ ；模型精度等级为合格时，则 $p > 0.8$ ， $C < 0.5$ ；模型精度等级为勉强合格时，则 $p > 0.7$ ， $C < 0.65$ ；模型精度等级为不合格时，则 $p \leq 0.7$ ， $C \geq 0.65$ 时，这种情况下预测模型需要修正，直到预测精度等级为合格后再预测。

3.3. 成都市老年人口预测

我们利用 Matlab (2014a)编程软件进行预测结果和精度检验，实际值与普通 GM(1,1)模型、优化 GM(1,1)模型对比如下表 3，具体预测值和精度检验如下表 4 和表 5。

由表 3 我们可知，优化的 GM(1,1)模型比普通 GM(1,1)模型在对实际值进行拟合的时候，很明显优化的模型拟合结果要好，说明优化是有提高精度的作用。

由表 4 我们可知，成都市的老年人口是呈持续增长状态的，由预测结果可知，2022 年的老年人口达到了 434 万人，而 2016 年才 300 万人，6 年时间增长了近 130 万人，增长速度之快，规模之大，值得我们关注老年人口，可知这预测结果是具有重要研究意义的。

表 5 是优化的 GM(1,1)模型预测的精度检验，包括普通 GM(1,1)模型和五次新陈代谢模型的检验。根据模型精度检验标准，结合精度检验结果可知，普通 GM(1,1)模型平均相对残差值为 0.0197，五次优化新陈代谢模型的平均相对残差最大值为 0.0167，当 $\alpha = 0.02$ 时，普通 GM(1,1)模型和优化 GM(1,1)新陈代

Table 2. The grade standard of model predictive accuracy test

表 2. 模型预测精度检验等级标准

检验指标	较好	合格	勉强合格	不合格
P	>0.95	>0.8	>0.7	≤ 0.7
C	<0.35	<0.5	<0.65	≥ 0.65

Table 3. The actual value of the population aging in Chengdu in 2011~2016: Comparison of general model and optimization model

表 3. 2011~2016 年成都市老年人口实际值：普通模型和优化模型对比表(单位：万人)

年份	2011	2012	2013	2014	2015	2016
实际值	214	223	236	249	260	300
GM(1,1)模型	211	225	239	254	271	287
优化模型	212	222	235	250	269	292

Table 4. The prediction value of population aging in Chengdu in 2017-2022 (unit: ten thousand people)

表 4. 2017~2022 年成都市老年人口预测值(单位：万人)

年份	2017	2018	2019	2020	2021	2022
优化模型	307	329	353	379	405	434

Table 5. Accuracy test
表 5. 精度检验表

模型精度	平均相对残差	关联度 R	小误差概率 P	后验方差比 C
普通 GM(1,1)模型	0.0197	61.8106%	1	0.201136
一次优化新陈代谢	0.0167	66.2660%	1	0.170145
二次优化新陈代谢	0.0140	69.6947%	1	0.143002
三次优化新陈代谢	0.0131	70.6008%	1	0.126634
四次优化新陈代谢	0.0125	71.8244%	1	0.115242
五次优化新陈代谢	0.0120	72.7630%	1	0.106322

谢模型的平均相对残差均小于 0.02, 由此我们可得无论是普通 GM(1,1)模型还是优化 GM(1,1)新陈代谢模型均为残差优秀模型; 在分辨率 $\lambda=0.5$ 时, 关联度 R 均大于 0.6, 说明普通 GM(1,1)模型和优化 GM(1,1)新陈代谢模型对原始数列的模拟度均为满意, 都通过了关联度检验; 小误差概率 P 均为 1, 且后验方差比 C 均小于 0.35, 表明普通 GM(1,1)模型和优化 GM(1,1)新陈代谢模型预测精度为较好, 模型对成都市老年人口进行了较高精度的预测, 预测结果适宜用于养老产业相关方面的分析。

4. 结束语

由上述实例分析可知, GM(1,1)预测模型的数据量要求小, 精度较高, 具有较强的实用性和有效性, 是比较理想的预测方法。对于开放性、非线性的复杂系统, GM(1,1)预测模型能够从整体出发, 对外延不确定性系统变化进行动态的科学模拟与仿真。因此, GM(1,1)预测模型适用于非线性系统的非唯一性预测拟合, 其预测结果可以为人口老龄化提供科学依据。

参考文献 (References)

- [1] 朱伟. 灰色预测模型及其应用[J]. 数学杂志, 1998, 18: 133-135.
- [2] 刘思峰, 党耀国, 方志耕, 等. 灰色系统理论及其应用[M]. 北京: 科学出版社, 2010.
- [3] 周瑞平. GM(1,1)模型灰色预测法预测城市人口规模[J]. 内蒙古师范大学学报, 2005, 34(1): 81-83.
- [4] 罗党, 刘思峰, 党耀国. 灰色模型 GM(1,1)优化[J]. 中国工程科学, 2003, 5(8): 50-53.
- [5] 王宁, 张爽, 曾庆均. 基于新陈代谢 GM(1,1)模型的重庆市人口老龄化预测研究[J]. 西北人口, 2017, 1(38): 66-70.
- [6] 王宁, 张爽. 基于 GM(1,1)模型的重庆市五大功能区人口老龄化预测[J]. 中国商论, 2016(20): 170-171.
- [7] 金浩, 张璐, 王志双. 基于组合预测的公路货运量需求预测实证研究——以唐山市为例[J]. 统计与管理, 2016(1): 52-55.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-2251, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: sa@hanspub.org