

Analysis on Guangdong Province Fertility Model Based on the Gray Prediction Method

Minhua Zhang, Xiner Lin, Liufang Wang, Meiting Huang

College of Science, South China Agricultural University, Guangzhou
Email: zmhwinnie2009wy@163.com

Received: Sep. 20th, 2014; revised: Oct. 23rd, 2014; accepted: Nov. 5th, 2014

Copyright © 2014 by authors 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

Abstract

As “China’s first populous province”, the movement in fertility rate in Guangdong has a close relationship with the province’s economic development and social environment in the future. Based on the statistical data, this article takes fertility for unilateral analysis using the statistical theory and methods, and analyzes its changing trend, finally makes a discussion about the trends.

Keywords

Fertility, Gray Prediction, Data Analysis, Model Checking

基于灰色预测方法的广东省人口生育率模型的研究

张敏华, 林馨儿, 王柳方, 黄美婷

华南农业大学理学院, 广州
Email: zmhwinnie2009wy@163.com

收稿日期: 2014年9月20日; 修回日期: 2014年10月23日; 录用日期: 2014年11月5日

摘要

作为“中国第一人口大省”的广东, 人口生育率的变动趋势与该省未来的经济发展和社会环境问题等都

有着密切的关系。本文基于统计数据,使用统计学的理论和方法,对生育率进行单方面的分析,并对其进行变动趋势分析,最后就该趋势进行研究和探讨。

关键词

生育率, 灰色预测, 数据分析, 模型检验

1. 引言

随着 2013 年中共十八大的召开,“二胎”新政策已引起了社会的广泛关注。广东省低生育率带来的老龄化问题,对家庭、对文化、对城市和国家经济社会协调和可持续发展的影响,都已经成为了迫在眉睫的民生问题。如何应对低生育率的人口形势而采取有效措施,是国家和相关部门亟需尽快解决的问题。探究和分析广东省人口生育率的变动趋势,对于是否实施“二胎”新政策以及因此而产生的经济、社会上的影响,是十分有必要的。传统的关于人口预测的相关研究大多从人口学和生物学的角度进行预测和分析,本文结合了统计学和应用数学的方法,采用数学建模中的灰色预测方法,通过建立数学模型,更科学客观的探讨分析了生育率的发展和变动,具有一定的创造性与现实意义。

2. 研究背景

早在 2011 年已有中国陷入超低生育率相关报道的出现。随着新中国的成立,在 1950~2100 年的 150 年间,中国总人口的运行轨迹就如同过山车。人口学者的主流预测是,在现行生育政策不变的前提下,中国的总人口峰值将远低于 15 亿,很可能低于 14.5 亿;同时,人口高峰发生的时间会在十多年后,此后将较快下降。即便按照乐观的估计,到 2100 年中国总人口规模也将缩小至 7.5 亿以内[1]。据此推测的生育率已经距离日本不远,此数据的潜在含义是近 15 年来的出生人口数量非常地少,意味着生育率非常低。人口问题不单纯是一个数字问题,它包含有更为广泛的社会经济、文化意义。由于几十年的计划生育政策的强制实施,由于生育率在社会经济发展下自动降低的社会规律,由于中华民族生育文化的长期压力与根深蒂固的矛盾,由于工业化、城市化的迅速推进和经济发展的不平衡,中国超低生育率过早的来临了,人口结构比例愈趋不合理。社会上关于“超低生育率”的关注已经进入白热化的阶段,各种言论一夜之间甚嚣尘上,计划生育政策更是成为众矢之的。好不容易平复了一段时间,自从 2013 年《中共中央关于全面深化改革若干重大问题的决定》关于启动实施一方为独生子女的夫妇可生育两个孩子的政策的提出,很多家庭都陷入了欢喜或忧愁当中[2]。因此,通过具体的实证分析“二胎”政策对未来社会、经济上的影响,具有极高的现实意义。

3. 研究思路与内容

建立广东省生育率模型。方法一是用现有的各年生育率(取 2006 年~2011 年)数据,通过灰色预测模型预测出未来五年广东省人口生育率;方法二则用已知的每年的出生人数及已知的年人均(以上数据取 2006 年~2011 年),使用灰色预测模型预测未来五年广东省的出生人数及年人均数,通过人口出生率(CBR)计算公式:出生率 = 年出生人数/年平均人数*1000‰,计算出对应年份的出生率。从统计方式来看,前者是基于统计数据直接得到的生育率数据,而后者是基于两项人口数据的联系计算得出的粗生育率。通过两种同种模型不同方法的预测,从理论以及对模型检验的分析,对比两种方法的可取性,最后判断模型的稳定性和可靠性。

4. 相关概念解释

4.1. 生育率

生育率是指一个群体中平均每个女性在其一生中当中所生育的子女数量，更替水平生育率为 2.1。

4.2. 出生率

出生率也称粗出生率，指在一定时期内(通常为一年)平均每千人所出生的人数的比率，一般用千分率表示。

4.3. 低生育水平

低生育水平是相对而言的。从我国的基本国情出发，在当前的经济条件下，总和生育率在更替水平与 1.8 之间，称为低生育水平。

5. 人口变动概况

5.1. 基于方法一的生育率概况

据《广东统计年鉴》调查数据显示，广东省 2006~2011 年人口生育率如表 1：

广东省人口生育率在 2006 年为 11.69，2007 年有小幅上升，2008 年以后开始下降，2011 年大幅下落，下降为最低点 10.45。

5.2. 基于方法二的出生率概况

广东省出生人口数在 2006 年为 108.96 万人，2007 年有小幅上升，2008 年以后上升幅度较大，2010 年以后大幅下落，2011 年下降为最低点 109.44 万人(表 2)。

广东省年人均数自 2006 年开始逐步稳健上升，趋势逐渐变缓。

6. 模型的建立与求解

6.1. 模型的建立

设原始序列为

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\},$$

Table 1. The fertility rate in Guangzhou province for each year

表 1. 广东省各年生育率

年份	2006	2007	2008	2009	2010	2011
出生率(%)	11.69	11.73	11.46	11.29	11.18	10.45

Table 2. The birth rate in Guangzhou province

表 2. 广东省人口出生率

年份	2006	2007	2008	2009	2010	2011
出生人口(万人)	108.96	112	112	113	115	109.44
年人均数(万人)	8048.71	8156.05	8267.09	8365.98	8521.55	8637.19
出生率(%)	13.53757	13.73214	13.54769	13.50708	13.49520	12.67079

首先对原始数据列中的数据进行处理, 利用累计生成(AGO)方法得到新数据列, 称生成序列

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$$

为原始序列的一次累加生成, 其中

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i) = x^{(1)}(k-1) + x^{(0)}(k), \quad k=1, 2, \dots, n$$

然后建立 GM(1,1)模型。称序列

$$Z^{(1)} = \{z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n)\}$$

为 $X^{(1)}$ 的紧邻均值生成序列, 其中

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1).$$

定义 GM(1,1)灰微分方程模型为

$$x^{(0)}(k) + az^{(1)}(k) = b,$$

式中 a 为反映 $\hat{x}^{(1)}$ 和 $\hat{x}^{(0)}$ 的发展态势的发展系数, b 为灰色作用量[3]。设 $\hat{\alpha} = (a, b)^T$ 为待估参数向量, 则灰微分方程的最小二乘估计参数列满足

$$\hat{\alpha} = (B^T B)^{-1} B^T Y_n,$$

其中

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}, \quad Y_n = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}.$$

时间响应序列为

$$\hat{x}^{(1)}(k+1) = \left[x^{(1)}(0) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a}, \quad k=1, 2, \dots, n-1$$

从而可得预测值

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$$

GM(1,1)模型的白化型为

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b.$$

6.2. 模型的求解

由于这里每个模型都只考虑了一个变量, 所以均采用 GM(1,1)模型。

以方法一中的生育率数据为例。设时间序列为

$$\begin{aligned} X^{(0)} &= \{x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), x^{(0)}(4), x^{(0)}(5), x^{(0)}(6)\} \\ &= \{11.69, 11.73, 11.46, 11.29, 11.18, 10.45\}. \end{aligned}$$

第一步。对 $X^{(0)}$ 作 1-AGO, 得

$$\begin{aligned} X^{(1)} &= \{x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), x^{(1)}(4), x^{(1)}(5), x^{(1)}(6)\} \\ &= \{11.69, 23.42, 34.88, 46.17, 57.35, 67.80\}. \end{aligned}$$

第二步。确定数据矩阵 B , Y , 得

$$B = \begin{bmatrix} -17.555 & 1 \\ -29.15 & 1 \\ -40.525 & 1 \\ -51.76 & 1 \\ -62.575 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} 11.73 \\ 11.46 \\ 11.29 \\ 11.18 \\ 10.45 \end{bmatrix}.$$

第三步。对参数列 $\hat{\alpha} = (a, b)^T$ 进行最小二乘估计, 得

$$\hat{\alpha} = (B^T B)^{-1} B^T Y = \begin{bmatrix} 0.0251 \\ 12.2343 \end{bmatrix}.$$

第四步。确定模型及时间响应式为

$$\begin{aligned} \frac{dx^{(1)}}{dt} + 0.0251x^{(1)} &= 12.2343 \\ \hat{x}^{(1)}(k+1) &= \left[x^{(0)}(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a} = -475.73231e^{-0.0251k} + 487.42231. \end{aligned}$$

第五步。由此得 $X^{(1)}$ 的模拟值

$$\hat{X}^{(1)} = \{11.69, 23.4821, 34.9817, 46.1962, 57.1326, 67.7978\}.$$

第六步。还原求出 $X^{(0)}$ 的模拟值

$$\hat{X}^{(0)} = \{11.69, 11.7921, 11.4996, 11.2145, 10.9364, 10.6652\}.$$

第七步。检验误差。记 $\varepsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$ 为残差, $\Delta(k) = \frac{|\varepsilon(k)|}{x^{(0)}(k)}$ 为相对误差[4](表 3)。

平均相对误差

$$\Delta = \frac{1}{6} \sum_{k=1}^6 \Delta(k) = 0.00965 = 0.965\%,$$

Table 3. Table on error inspection
表 3. 误差检验表

序号	实际数据 $x^{(0)}(k)$	模拟数据 $\hat{x}^{(0)}(k)$	残差 $\varepsilon(k)$	相对误差 $\Delta(k)$
1	11.69	11.6900	0	0
2	11.73	11.7921	0.0621	0.0053
3	11.46	11.4996	0.0396	0.0035
4	11.29	11.2145	-0.0755	0.0067
5	11.18	10.9364	-0.2436	0.0218
6	10.45	10.6652	0.2152	0.0206

Table 4. The result of the grey prediction method

表 4. 灰色预测方法所得到的结果

	模型	初始值		模拟值		预测值		相对误差		结论
方法一	$\frac{dx^{(1)}}{dt} + 0.0251x^{(1)} = 12.2343$	11.69	11.73	11.6900	11.7921	10.4007	10.1428	0	0.0053	可进行中长期预测
		11.46	11.29	11.4996	11.2145	9.8913	9.6460	0.0035	0.0067	
		11.18	10.45	10.9364	10.6652	9.4068		0.0218	0.0206	
数据一	$\frac{dx^{(1)}}{dt} + 0.0019x^{(1)} = 113.0179$	108.96	112	108.9600	112.7086	111.6590	111.4503	0	0.0063	可进行中长期预测
		112	113	112.4979	112.2876	111.2420	111.0340	0.0044	0.0063	
		115	109.44	112.0777	111.8682	110.8265		0.0254	0.0222	
数据二	$\frac{dx^{(1)}}{dt} + 0.0145x^{(1)} = 7971.9$	8048.71	8156.05	8048.7	8147.7	8760.9	8889.0	0	0.0010	可进行中长期预测
		8267.09	8365.98	8266.8	8387.7	9018.9	9150.8	0.0000	0.0026	
		8521.55	8637.19	8510.3	8634.7	9284.6		0.0013	0.0003	

预测值的模拟精度基本满足要求，因此可以用来预测[5]。

由 $\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$ 可以预测出未来五年广东省人口生育率为

$$\begin{aligned} \hat{X}^{(0)} &= \{\hat{x}^{(0)}(7), \hat{x}^{(0)}(8), \hat{x}^{(0)}(9), \hat{x}^{(0)}(10), \hat{x}^{(0)}(11)\} \\ &= \{10.4007, 10.1428, 9.8913, 9.6460, 9.4068\}. \end{aligned}$$

方法二的模型建立与求解同理可得。

7. 分析预测

基于以上所提供的数据资料，使用 MATLAB 软件通过灰色预测方法可以得到如表 4 所示结果：

从方法一可以看出，模型通过了检验，具有一定的稳定性和可靠性，并且预测出广东省未来五年的人口生育率分别为：10.4007、10.1428、9.8913、9.6460、9.4068。同时，方法二中的模型也都通过了检验，通过人口出生率(CBR)计算公式：出生率 = 年出生人数/年平均人数*1000‰，预测出广东省未来五年的出生率分别为：12.74525、12.53798、12.33404、12.13338、11.93701。

由此可得，广东省人口生育率日渐下降，在未来五年将维持总体下滑的趋势。

8. 结语

人口问题不仅是民生问题，更与国家与社会的各个方面息息相关。在现今的科学依据和形势预测上，如何改善中国低生育率的状况，是一个值得探讨的课题[6]。要解决这一问题，不仅要靠相关政策的出台，最重要的还是人们自己的观念和态度，只有全社会一起共同执行，互相监督，人口工作才能顺利完成。

基金项目

华南农业大学大学生创新训练项目(201310564038)。

参考文献 (References)

- [1] 苏岭, 姚永泳 (2011) 中国陷入超低生育率陷阱. 南方周末.
- [2] 张艳红 (2013) “单独二胎”政策出台 有人欢喜有人愁. 钱江晚报.
- [3] 房少梅 (2014) 数学建模理论、方法及应用. 科学出版社, 北京.
- [4] 王庚, 王敏生 (2008) 现代数学建模方法. 科学出版社, 北京.
- [5] 严喜祖, 宋中民, 毕春加 (2009) 数学建模及其实验. 科学出版社, 北京.
- [6] 党耀国, 刘思峰, 王正新, 林益 (2009) 灰色预测与决策模型研究. 科学出版社, 北京.